AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LIII, Number 2

APRIL, 1931

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

Upon a theory of infinite systems of non-linear implicit and differential equations. By AUREL WINTNER, ... 241
Note on the numerical value of a particular mass ratio in the restricted problem of three bodies. By JENNY E. ROSENTHAL, . 258
Upon the existence and non-existence of isoenergetic periodic perturbations of the undisturbed circular motions in the restricted problem of three bodies. By MONROE MARTIN, . . 259
Note on the constants of the disturbing function. By K. P. WILLIAMS, 274
Conductors in an electromagnetic field $(\boldsymbol{E}^0e^{pt}, \boldsymbol{H}^0e^{pt})$. By F. H. MURRAY, 275
The gravitational field of a body with rotational symmetry in Einstein's theory of gravitation. By P. Y. CHOU, ... 289
On the non-existence of curves of order 8 with 16 cusps. By OSCAR ZARISKI, ... 309
Construction of pencils of equianharmonic cubics. By JACOB YERUSHALMY, ... 319
The planar imprimitive group of order 216. By JOHN ROGERS MUSSELMAN, ... 333
A prepared system for two quadratics in six variables. By J. WILLIAMSON, ... 343
A variety representing pairs of points of space. By F. R. SHARPE, . 358
On semi-metric spaces. By WALLACE ALVIN WILSON, ... 361
Concerning hereditarily locally connected continua. By G. T. WHYBURN, 374
Generalizations of biconnected sets. By P. M. SWINGLE, . . . 385
Some metric properties of descriptive planes. By J. L. DORROH, . 401
Concerning metric collections of continua. By J. H. ROBERTS, . . 422
The cyclic and higher connectivity of locally connected spaces. By G. T. WHYBURN, ... 427
Generalized Green's matrices for compatible systems of differential equations. By WILLIAM T. REID, ... 443
A Cremona group of order thirty-two of cubic transformations in three-dimensional space. By ETHEL ISABEL MOODY, . . 460
On the capacity of sets of Cantor type. By OLIVER D. KELLOGG, . 475
Note on fractional operators and the theory of composition. By LEONARD M. BLUMENTHAL, ... 483
Potentials of general masses in single and double layers. The relative boundary value problems. By G. C. EVANS and E. R. C. MILES, ... 493

THE AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 25 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. COHEN at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918

PRINTED BY THE J. H. FURST COMPANY
BALTIMORE, MD.

UPON A THEORY OF INFINITE SYSTEMS OF NON-LINEAR IMPLICIT AND DIFFERENTIAL EQUATIONS.

By Aurel Wintner.

Introduction. There has arisen, in recent years, a large literature upon infinite systems of non-linear implicit and differential equations.[1] The methods which are employed have become classical in the treatment of finite systems. However these methods (employed in the treatment of finite systems) may be extended only to infinite systems of very special type, restricted by inequalities, which for many reasons are not fulfilled in possible applications of the method of infinitely many variables to various classical problems in analysis. Thus the theorems which have been proved by me a few years ago[2] are the only ones to my knowledge which have found an application to concrete problems, in particular to the problems of Celestial Mechanics or the Calculus of Variations. I shall give here a comprehensive review of these questions without giving any essential extensions of my results as set forth in my previous papers and without assuming any previous knowledge of the theory of infinitely many variables. The applications will be excluded and the reader referred in this connection to some earlier papers appearing in the *Mathematische Annalen* and in the *Mathematische Zeitschrift.* A characteristic application can be found in the paper of Martin appearing in this number of the Journal.[3]

The infinite system which I shall treat is composed exclusively of power series (and not of more general functions) of the infinite sequence of variables (as is indeed always the case in concrete applications).

My problem was to introduce a method of treatment in which the power series are not subjected to inequalities which would not be fulfilled in concrete applications but which would be necessary in the classical modes of treatment. On page 252 there will be given a number of examples, of most important character, which will serve to illustrate why the usual methods *must* fail.

Inasmuch as the nature of the problems appears most clearly in its application to the complex space of Hilbert, the present paper will be restricted in its treatment to this space. It is of course possible by using the principles of General Analysis[4] to further extend the methods here set forth. There is no difficulty in using other spaces[5] than that of Hilbert. In the articles cited above also spaces other than that of Hilbert are treated; cf. the above mentioned paper of Martin.

The systems to which the existence theorems of this paper apply are simply the direct generalizations of the most general bounded (beschränkt) systems of linear equations of Hilbert, or the corresponding linear differential systems, to non-linear systems. The notion of a regular power series in infinitely many variables, which enters in the existence theorems treated in this paper, will be so defined that, in the special cases of linear and quadratic forms, the regular power series are nothing other than the bounded linear and quadratic forms of Hilbert (the coefficients of which are not necessarily real).[6]

1. *Regular power series.* By a power series in infinitely many variables $z_1, z_2, \cdots$ will be understood formally an expression

$$(1) \qquad \Phi(z_1, z_2, \cdots) = \sum_{n=0}^{\infty} \Phi^{(n)}(z_1, z_2, \cdots),$$

in which

$$(2) \qquad \Phi^{(n)}(z_1, z_2, \cdots) = \sum_{\nu_1=1}^{\infty} \sum_{\nu_2=1}^{\infty} \cdots \sum_{\nu_n=1}^{\infty} c^{(n)}{}_{\nu_1\nu_2\ldots\nu_n} z_{\nu_1} z_{\nu_2} \cdots z_{\nu_n},$$

that is $\Phi^{(n)}$ is a form of degree n. It is clear that any such form (in which the coefficients and variables may be complex) can, without loss of generality, be written so that

$$(3) \qquad c^{(n)}{}_{\nu_1\nu_2\ldots\nu_n} = c^{(n)}{}_{\nu_2\nu_1\ldots\nu_n} = \cdots = c^{(n)}{}_{\nu_n\nu_{n-1}\ldots\nu_1},$$

and this will accordingly, in the following, always be assumed to have been done.

We now define what is meant by the convergence of the power series

$$(4) \qquad \Phi(z_1, z_2, \cdots, z_m, z_{m+1}, z_{m+2}, \cdots) = \sum_{n=0}^{\infty} \sum_{\nu_1=1}^{\infty} \sum_{\nu_2=1}^{\infty} \cdots \sum_{\nu_n=1}^{\infty} c^{(n)}{}_{\nu_2\nu_1\ldots\nu_n} z_{\nu_1} z_{\nu_2} \cdots z_{\nu_n}$$

[and therefore also in the special case of a form (2)]. We put all the variables, with the exception of the first m, equal to zero and obtain a power series in the m variables $z_1, z_2, \cdots z_m$, namely

$$(5m) \qquad \Phi_{[m]}(z_1, z_2, \cdots z_m) \equiv \Phi(z_1, z_2, \cdots, z_m, 0, 0, \cdots) = \sum_{n=0}^{\infty} \sum_{\nu_1=1}^{m} \sum_{\nu_2=1}^{m} \cdots \sum_{\nu_n=1}^{m} c^{(n)}{}_{\nu_1\nu_2\ldots\nu_n} z_{\nu_1} z_{\nu_2} \cdots z_{\nu_n}.$$

This power series, arising from (4), will be called the m-th section (Abschnitt) of (4). We say that the power series (4) converges at a point

$$(6) \qquad z_1, z_2, \cdots$$

of the space of infinitely many dimensions if first: At the point (6) every

section $(5m)$ converges (in the usual sense of the theory of m-fold series) for $m = 1, 2, \cdots$ (such a restriction is obviously by its very nature fulfilled in the case of forms or polynomials in infinitely many variables) and if secondly: At the point (6) the limit

$$(7) \qquad \lim_{m\to\infty} \Phi_{[m]}(z_1, z_2, \cdots z_m)$$

exists. The sum of the series (4) at the point (6) is defined as the value of this limit.

The power series

$$(4') \qquad \sum_{n=0}^{\infty} \sum_{\nu_1=1}^{\infty} \sum_{\nu_2=1}^{\infty} \cdots \sum_{\nu_n=1}^{\infty} | c^{(n)}{}_{\nu_1\nu_2\ldots\nu_n} | z_{\nu_1} z_{\nu_2} \cdots z_{\nu_n},$$

which arises from (4) when its coefficients are replaced by their absolute values, we shall call the best majorant of (4) and will be designated by $\bar{\Phi}$. If the series (4′) converges at the point

$$(6') \qquad | z_1 |, \, | z_2 |, \cdots,$$

the series (4) is said to be absolutely convergent at the point (6).

With respect to a given sequence of numbers

$$(8) \qquad z_1^{(0)}, z_2^{(0)}, \cdots$$

the symbol

$$(9) \qquad \sigma_r\{z_k^{(0)}\}$$

will be understood to designate the following domain of the infinitely many independent variables (6):

$$(10) \qquad \sum_{k=1}^{\infty} | z_k - z_k^{(0)} |^2 < r^2.$$

Accordingly (9) is the interior of the complex Hilbert sphere, with radius r, taken about the point (8) as center.

In particular we shall write briefly

$$(11) \qquad \sigma_r = \sigma_r\{0\},$$

so that

$$(12) \qquad \sigma_r: \qquad \sum_{k=1}^{\infty} | z_k |^2 < r^2.$$

We say [7] that the power series (4) is regular in the domain (12) if the following three conditions are fulfilled: First, the m-th section $(5m)$ of (4) shall, for any fixed m, in the sense of the theory of analytic functions of m variables be regular in the m-th "section" of σ_r, that is, in the domain

$$(13m) \qquad \sum_{k=1}^{m} | z_k |^2 < r^2.$$

Second, there exists at every point of (12) the limit defined by (7); the series (4) is then, according to the above definition, convergent. Third, there exists for every positive number ϵ a positive number $M_{r-\epsilon}$ which is independent of m and such that the absolute value of the power series ($5m$) in the domain

$$(13m') \qquad \sum_{k=1}^{m} | z_k |^2 < (r - \epsilon)^2$$

is not greater than $M_{r-\epsilon}$; in other words the absolute values of the function (4) in the domain $\sigma_{r-\epsilon}$ is not greater than $M_{r-\epsilon}$. It follows from the first assumption that the power series ($5m$) converges absolutely in the domain ($13m$) and uniformly in the domain ($13m'$). It is clear that the second assumption is not a consequence of the first. In the special cases where the power series (4) is a linear or a quadratic form Hellinger and Toeplitz [8] have shown that the third assumption is a consequence of the first two. However it is easy to show that for the general power series which contain terms of infinite degree the third assumption is independent of the first two. The power series (4) will be said to be an integral function if it is regular in a σ_r with arbitrarily large r.

The upper bound of $| \Phi(z_1, z_2, \cdots) |$ in the domain σ_ρ will be designated by $[\Phi]_\rho$. In the domain $0 \leqq \rho < r$ the upper bound $[\Phi]_\rho$ is a positive monotone function of ρ which, while it is bounded in the interval $0 \leqq \rho < r - \epsilon$ for any given value of ϵ, can become infinite if $\epsilon \to 0$ (Φ being regular in σ_r). The function Φ can be an integral function and nevertheless the series (4) not be uniformly convergent in any domain σ_δ (with arbitrarily small δ). The simplest example of an integral function of this kind is

$$(14) \qquad \Phi = \sum_{k=1}^{\infty} z_k^2.$$

Moreover it is possible that Φ, for example even in the special case of the quadratic form

$$(15) \qquad \Phi = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} c_{ij} z_i z_j,$$

is an integral function and nevertheless the series (4) is not absolutely convergent in a domain σ_δ regardless of how small δ be chosen. We shall return to this point later (cf. page 248). In particular (4) can accordingly be an integral function without the necessary existence of a domain σ_δ in which

the best majorant is regular. If now not only Φ but also $\bar{\Phi}$ is regular in a domain σ_r we shall say Φ is absolutely regular in σ_r. A regular function which is a form is, because of its homogeneity, obviously an integral function. Inasmuch as the sum of a finite number of power series, regular in σ_r, is a power series regular in σ_r, we may say more generally that the sum of a finite number of regular forms, i. e. a regular polynomial, is likewise an integral function.

For a given function Φ regular in σ_r it is possible in the neighborhood of a given point (8) of σ_r to order formally the function according to increasing powers of the arguments $z_k - z_k^{(0)}$. One thus obtains a power series in the arguments $z_k - z_k^{(0)}$ which in the domain

$$(10') \qquad \sum_{k=1}^{\infty} | z_k - z_k^{(0)} |^2 < r'^2, \quad r'^2 = r^2 - \sum_{k=1}^{\infty} | z_k^{(0)} |^2$$

is[9] regular and for corresponding points possesses the same sum as Φ. This theorem will be called the neighborhood theorem and implies among other things that in the concept of neighborhood associated with the metric $(\sum_{i=1}^{\infty} | u_i - v_i |^2)^{1/2}$ there exists no isolated regular point. Nevertheless (because of the nature of the topology of $\sum_{i=1}^{\infty} | x_i |^2 < \delta$) it can not be expected that there exist, in the theory of functions of infinitely many variables, far reaching analogies with the theory of analytic functions of a finite number of variables. As an example, the theorem of Poincaré-Volterra, based on the theorem of Heine-Borel, is no longer true, for Hilbert gives the example[7]

$$\Phi = \sum_{\nu=0}^{\infty} \sum_{i=1}^{\infty} \frac{(-1)^\nu}{2^i} \binom{1/2}{\nu} z^\nu = \sum_{i=1}^{\infty} \frac{(1 - z_i)^{1/2}}{2^i}$$

of an obviously regular (also in our sense) function for which the set of branches has the power of the continuum.

Given the sequence of points $\{z_k^{(0)}\}$, $\{z_k^{(1)}\}$, $\{z_k^{(2)}\}, \cdots$ in the domain σ_r in which the function Φ is defined. The function Φ will be said to be continuous at the point $\{z_k^{(0)}\}$ if for every sequence in which

$$(16) \qquad \lim_{\nu\to\infty} z_1^{(\nu)} = z_1^{(0)}, \quad \lim_{\nu\to\infty} z_2^{(\nu)} = z_2^{(0)}, \quad \lim_{\nu\to\infty} z_3^{(\nu)} = z_3^{(0)}, \cdots$$

we always have

$$(17) \qquad \lim_{\nu\to\infty} \Phi(z_1^{(\nu)}, z_2^{(\nu)}, z_3^{(\nu)}, \cdots) = \Phi(z_1^{(0)}, z_2^{(0)}, z_3^{(0)}, \cdots).$$

There exist regular, and indeed integral functions, which are continuous at no point. The simplest example is the afore-mentioned function (14).

To be sure it may be shown [10] that if Φ is regular in σ_r and (8) is a point of σ_r then (17) is valid provided that the sequence (16) converges strongly to $\{z_k^{(0)}\}$, that is if not (16), but also

$$\lim_{\nu\to\infty} \sum_{k=1}^{\infty} | z_k^{(\nu)} - z_k^{(0)} |^2 = 0 \tag{18}$$

is supposed.

If we designate by $\Phi_{z_k}(z_1, z_2, \cdots)$ the power series arising from $\Phi(z_1, z_2, \cdots)$ by term by term partial differentiation with respect to z_k, it can be shown [11] that if Φ is regular in σ_r, then also Φ_{z_k} is regular in σ_r. Moreover we have [12]

$$\sum_{k=1}^{\infty} \left| \frac{\partial\Phi(z_1, z_2, \cdots)}{\partial z_k} \right|^2 \leqq \left(\frac{[\Phi]_{r-\epsilon}}{r-\epsilon} \right)^2 = \text{const.} \quad \text{in } \sigma_{r-\epsilon} \quad (\epsilon > 0). \tag{19}$$

This inequality will be designated as the gradient inequality.

2. *The bounded forms of Hilbert considered as regular power series.* We introduce now a new complex manifold $\boldsymbol{E}$

$$\boldsymbol{E}\colon \sum_{i=1}^{\infty} | \zeta_i |^2 = 1 \tag{20}$$

in the independent variables ζ_i. The linear form

$$\sum_{i=1}^{\infty} a_i \zeta_i \tag{21}$$

is bounded in the sense of Hilbert if and only if the series [6]

$$\sum_{i=1}^{\infty} | a_i |^2 \tag{22}$$

converges. It follows from the inequality of Schwarz that the absolute value of (21) is not greater than the square root of the expression (22). That the expression (21) actually attains its maximum in $\boldsymbol{E}$ follows readily if we place in (21)

$$\zeta_i = \frac{\bar{a}_i}{\left(\sum_{k=1}^{\infty} | a_k |^2 \right)^{1/2}} . \tag{23}$$

Accordingly the linear form

$$\sum_{i=1}^{\infty} c_i z_i \tag{24}$$

is then and only then bounded in the sense of Hilbert if, considered as a power series in infinitely many variables, it is a regular function (and accordingly

an integral function) in our sense. Accordingly if (24) is a regular function in σ_ρ we have (with the use of the notation introduced on page 244)

$$(25)\qquad [\sum_{k=1}^{\infty} c_k z_k]_\rho = \rho(\sum_{k=1}^{\infty} | c_k |^2)^{1/2}.$$

For brevity we write

$$(26)\qquad [\{c_k\}] = (\sum_{k=1}^{\infty} | c_k |^2)^{1/2}.$$

Since the convergence of the series (26) represents the necessary and sufficient condition for the regularity of (24) it follows that (24) and

$$(24')\qquad \sum_{i=1}^{\infty} | c_i | z_i$$

are simultaneously regular. Or more concisely, a regular *linear* form is always absolutely regular.

The matrix $\| a_{ij} \|$, which may be complex and need not fulfill any symmetry conditions, is then and only then bounded in the sense of Hilbert[6] if each of the linear forms $\sum_{j=1}^{\infty} a_{ij}\zeta_j$ $(i = 1, 2, \cdots)$ is bounded in the sense of Hilbert and if in addition there exists a positive number K so that in each point of $\boldsymbol{E}$

$$(27)\qquad \sum_{i=1}^{\infty} | \sum_{j=1}^{\infty} a_{ij}\zeta_j |^2 \leqq K.$$

In order that a matrix be bounded in the sense of Hilbert it is necessary[13] that there exist a constant L independent of m so that

$$(28)\qquad | \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\zeta_i\zeta_j | \leqq L$$

in every point of

$$(28')\qquad \sum_{k=1}^{m} | \zeta_k |^2 = 1.$$

It is furthermore necessary that the double series

$$(29)\qquad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij}\zeta_i\zeta_j$$

converge in every point of $\boldsymbol{E}$. The sum of the series is then, from (28), obviously $\leqq L$ in absolute value. The exact upper limit $\leqq K^{1/2}$ of

$$(27')\qquad (\sum_{i=1}^{\infty} | \sum_{j=1}^{\infty} a_{ij}\zeta_j |^2)^{1/2}$$

in the domain $\boldsymbol{E}$ will be designated by $[\ \| a_{ij} \|\ ]$. If $\| a_{ij} \|$ is bounded in

the sense of Hilbert it is known that the transposed matrix is also bounded in the sense of Hilbert and that $[\| a_{ij} \|] = [\| a_{ji} \|]$.

We now assume that the matrix $\| a_{ij} \|$ is symmetric ($a_{ij} = a_{ji}$) without necessarily being real. Hellinger and Toeplitz[6] have shown that in this case the existence of a positive number L, independent of m, so that (28) is valid in every point of the domain (28′) is not only necessary, but also sufficient in order that the matrix $\| a_{ij} \|$ should be bounded in the sense of Hilbert. Since we have agreed (3) to write every quadratic form in the symmetric form

$$(30) \qquad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} c_{ij} z_i z_j \qquad (c_{ij} = c_{ji})$$

(in which the coefficients may be complex) it follows from the foregoing that every form (30), that is, a not necessarily real, but symmetric matrix $\| c_{ij} \|$ is bounded in the sense of Hilbert if and only if the form (30), regarded as a power series in infinitely many variables, is a regular function (and therefore an integral function). It follows from the definitions (cf. page 244 and page 247) and from the inequality of Schwarz that

$$[\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} c_{ij} z_i z_j]_\rho \leqq \rho^2 [\, \| c_{ij} \| \,].$$

Toeplitz[14] has shown that there exist symmetric forms (30) which are bounded in the sense of Hilbert, although the best majorant form

$$(30') \qquad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} | c_{ij} | z_i z_j$$

is not bounded in the sense of Hilbert. The justification of the remarks made in connection with (15) accordingly follows from the above theorem of Hellinger and Toeplitz [the example given by Toeplitz is moreover real and not only bounded but also in the sense of Hilbert "completely continuous" (vollstetig), i. e. in the sense of the definition given in connection with (16), (17) continuous in every point of the Hilbert space].

3. *Bounded sequences and bounded matrices of regular power series.*

We will say that the vector $\{\Phi_i(z_1, z_2, \cdots)\}$ with infinitely many (not necessarily real) components

$$(31) \qquad \Phi_1, \Phi_2, \Phi_3, \cdots$$

is bounded in σ_r if the following three conditions are satisfied:

(a) The power series

$$(32) \qquad \Phi_i \equiv \Phi_i(z_1, z_2, \cdots); \qquad (i = 1, 2, \cdots)$$

are regular, in the sense defined above, in the domain σ_r.

(b) If $\{z_k\}$ is any arbitrarily chosen fixed point of σ_r then the linear form

$$\sum_{i=1}^{\infty} \Phi_i(z_1, z_2, \cdots)\zeta_i \tag{33}$$

is bounded in the sense of Hilbert, i. e. the number (cf. page 244)

$$[\{\Phi_i(z_1, z_2, \cdots)\}] = \sqrt{\sum_{i=1}^{\infty} |\Phi_i(z_1, z_2, \cdots)|^2}, \tag{34}$$

which depends upon $\{z_k\}$ is $< +\infty$ for any point in σ_r.

(c) The expression (34) is a bounded function of $\{z_k\}$ in the domain σ_r i. e. there exists a constant M so that for every point $\{z_k\}$ of the domain σ_r

$$[\{\Phi_i(z_1, z_2, \cdots)\}] \leqq M. \tag{34'}$$

The smallest M i. e. the upper limit of (34) in the domain σ_r will be designated by

$$[\{\Phi_i\}]_r. \tag{35}$$

In the special case when the vector components (31) are constants the above definition of a bounded vector is equivalent to Hilbert's definition of the boundedness of the linear form

$$\sum_{i=1}^{\infty} \Phi_i(0, 0, \cdots)\zeta_i. \tag{33'}$$

In an analogous manner we shall generalize the concept of boundedness, in the sense of Hilbert, for a matrix with constant elements, to a matrix whose elements are power series, as follows:

We shall say that the matrix

$$\| \Phi_{ij}(z_1, z_2, \cdots) \| \tag{36}$$

(which need neither be real nor satisfy any symmetry conditions whatsoever) is bounded in σ_r if the following three conditions are fulfilled:

(a) Each of the power series

$$\Phi_{ij}(z_1, z_2, \cdots); \qquad (i, j = 1, 2, \cdots) \tag{37}$$

is regular in σ_r.

(b) The matrix (36) is bounded in every point $\{z_k\}$ of σ_r, i. e. the number

$$[\| \Phi_{ij}(z_1, z_2, \cdots) \|], \tag{38}$$

defined on page 247, which depends upon $\{z_k\}$ remains finite for any point in σ_r.

(c) There exists a constant M so that (38) [in every point of σ_r] is not

greater than M; the smallest M [i. e. the upper limit of (38) in the domain σ_r] will be designated by

(39) $$[\| \Phi_{ij} \|]_r.$$

The following consistency theorems are now valid:

I. If the power series (4) is regular in σ_r and if ϵ designates an arbitrarily small positive number, the vector with components

(40) $$\Phi_i \equiv \Phi_{z_i} \equiv \partial\Phi/\partial z_i$$

is bounded in the domain $\sigma_{r-\epsilon}$ and satisfies the inequality

(41) $$[\{\Phi_{z_i}\}]_{r-\epsilon} \leqq \frac{[\Phi]_{r-\epsilon}}{r-\epsilon} \qquad \text{for every } \epsilon > 0.$$

II. If (31) is a bounded vector in the domain σ_r then the Jacobian matrix

(42) $$\| \partial\Phi_i/\partial z_k \|$$

is, for arbitrarily small $\epsilon > 0$, a bounded matrix in the domain $\sigma_{r-\epsilon}$.

III. If (4) is a regular power series in the domain σ_r the Hessian matrix

(43) $$\| \partial^2\Phi/\partial z_i \partial z_j \|$$

is, for arbitrarily small $\epsilon > 0$, a bounded matrix in the domain $\sigma_{r-\epsilon}$.

I follows immediately from the gradient inequality (19). II follows from I if we put $\Phi = \sum_{k=1}^{\infty} \Phi_k(z_1, z_2, \cdots)\zeta_k$ in I. III is a trivial consequence of I and II.

The multiplication theorems (Hilbertsche Faltungssätze) and related theorems of Hilbert upon matrices with constant elements can, by the usual proofs,[6] be extended to our more general case so that I shall not enter any further in these matters. However the following trivial theorem, which will be needed later, will now be formulated:

IV. Any vector (31), bounded in σ_r, is transformed by a matrix $\| a_{ij} \|$ into a vector

(44) $$\psi_1, \psi_2, \cdots; \qquad \psi_i = \sum_{j=1}^{\infty} a_{ij}\Phi_j(z_1, z_2, \cdots)$$

which is likewise bounded in σ_r.

In addition the following, easily proven, theorems are valid.[15]

V. If (31) is a bounded vector in σ_r [or if (36) is a bounded matrix in σ_r] and (8) is an arbitrarily chosen fixed point of σ_r and if one orders

every power series Φ_i of the vector (31) [or every power series Φ_{ij} of the matrix (36)] in the neighborhood of the point (8) formally according to increasing powers of the arguments $z_k - z_k^{(0)}$, one obtains a vector (or a matrix) which is bounded in the domain (10′) (cf. page 245) of the new variables $z_k - z_k^{(0)}$.

VI. The sum of two vectors (or matrices) bounded in σ_r is a vector (or matrix) bounded in σ_r.

This follows immediately with the use of the inequality of Schwarz.

4. *The existence theorems.* In the proof of existence theorems on infinite systems, the above consistency theorems permit us to take as a starting point in the proof a conveniently normalized form of the infinite system under consideration.

The fundamental existence theorem of the theory is the following:

EXISTENCE THEOREM I.[16] *If the vector* (31) *is bounded in σ_r the system*

$$dz_i/dt = \Phi_i(z_1, z_2, \cdots); \qquad (i = 1, 2, \cdots) \tag{45}$$

possesses in the circle

$$|t| < \frac{r}{2[\{\Phi_i\}]_r} \tag{46}$$

a holomorphic solution

$$z_i = z_i(t); \qquad (i = 1, 2, \cdots), \tag{47}$$

satisfying the initial conditions

$$z_i(0) = 0; \qquad (i = 1, 2, \cdots), \tag{48}$$

and this solution (47) *lies in the domain σ_r if t lies in the domain* (46).

[The uniqueness of the solution (47), which is more interesting than in the case of a finite number of variables is treated *loc. cit.*[2], III, p. 466-467].

It follows from the Consistency Theorem V that the normalization (48) can be effected without loss of generality. In addition there is no loss in generality in assuming that the functions Φ_i of (45) do not contain the independent variable t explicitly, inasmuch as one can in this case adjoin to (45) the differential equation

$$dz_0/dt = 1 \qquad [z_0(0) = 0], \tag{45′}$$

without thereby destroying the boundedness of the vector whose components are the right hand members of (45).

For the special case of a linear system of differential equations Hart[17] has demonstrated this existence theorem by a method which cannot be ex-

tended to the non-linear case.[18] The existence domain found by Hart in the special case of the linear systems is larger than that given in (46), valid for the above more general case. In fact the existence domain [19] found by Hart can be derived as a corollary from Existence Theorem I simply by analytical continuation in which one need only use the divergence of the harmonic series.[20]

Mention has been made, in the introduction, of the failure of the usual existence proofs under our general assumptions. We intend now to discuss this matter in greater detail. First of all, the method of Arzelà and related methods,[21] in which only the continuity [or even less restrictive assumptions [22]] of the Φ_i is assumed, are not applicable. This appears immediately from the known fact that the Hilbert sphere is neither separable nor compact. The methods employed under the condition of Lipschitz are also not applicable, as is readily perceived from the integral function (14) which is neither continuous nor satisfies the condition of Lipschitz. That the majorant method of Cauchy, which assumes the existence of a single majorant equation

$$dz/dt = \phi(z), \tag{49}$$

must also fail follows from the example

$$\Phi_i = z_i; \qquad (i = 1, 2, \cdots). \tag{50}$$

The vector (31) is, in the case of (50), bounded in any σ_r. The best common majorant, with one variable, of the infinitely many power series (50) is however

$$\sum_{i=1}^{\infty} z = z \sum_{i=1}^{\infty} 1, \tag{50'}$$

a divergent series, so that the best common majorant (49) does not exist. This type of infinite differential system, in which the variable z_i dominates the function Φ_i, is moreover a typically occurring case in the applications.

The method of Cauchy is a special majorant method, as it is based on the existence of a *single* (*common*) majorant differential equation. It will now be shown that the proof of Existence Theorem I lies beyond the capabilities of the majorant method even if one replaces each of the infinitely many differential equations by its best majorant, i. e. even if one works with an infinite majorant *system*. The best majorant system of (45) is [cf. (4′)]

$$dz_i/dt = \bar{\Phi}_i(z_1, z_2, \cdots). \tag{51}$$

One can now so choose the Φ_i that on one hand the vector (31) is bounded

in every domain σ_r, and consequently the system (45), according to Existence Theorem I possesses a holomorphic solution, satisfying the initial condition (48), while on the other hand the best majorant vector

$$\bar{\Phi}_1, \bar{\Phi}_2, \cdots \tag{52}$$

is not only not bounded in a sufficiently small domain σ_r but the best majorant system (51) of (45) cannot possess a holomorphic solution satisfying the initial conditions (48). [In what follows, for the sake of brevity, we shall demonstrate only the non-existence of a holomorphic solution of (51) satisfying (48)]. We infer accordingly that the best majorant of a solution is not to be confused with the solution of the best majorant system.

In order to give an example of a differential system (45) which fulfills the conditions of Existence Theorem I but which does not yield to the devices of *any* majorant method we put in (45) at first

$$\Phi_i \equiv a_i + \sum_{k=1}^{\infty} a_{ik} z_k \tag{53}$$

where the form

$$\sum_{i=2}^{\infty} \sum_{k=2}^{\infty} | a_{ik} | z_i z_k$$

is not bounded in the sense of Hilbert. The two bounded linear forms

$$\sum_{k=2}^{\infty} a_{1k} X_k, \qquad \sum_{k=2}^{\infty} a_k X_k$$

can be so chosen that [8]

$$\sum_{\nu=2}^{\infty} \sum_{\mu=2}^{\infty} | a_{\nu\mu} a_{1\nu} a_\mu | = + \infty.$$

If we now assume that

$$\sum_{i=2}^{\infty} \sum_{k=2}^{\infty} a_{ik} z_i z_k$$

is symmetrical and bounded in the sense of Hilbert (which contains no contradiction [14]) and if we place

$$a_1 = 0, \qquad a_{i1} = 0; \qquad\qquad (i = 1, 2, \cdots),$$

the premises of the Existence Theorem I are fulfilled since the two forms

$$\sum_{i=1}^{\infty} a_i z_i, \qquad \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} a_{ik} z_i z_k$$

are obviously bounded in the sense of Hilbert; while the best majorant system (51) of (45), (53), namely the system

$$dz_i/dt = | a_i | + \sum_{j=1}^{\infty} | a_{ij} | z_j ; \qquad (i = 1, 2, \cdots)$$

possesses no holomorphic solution

$$z_i(t) = \sum_{k=1}^{\infty} \gamma_{ik} t^k$$

since the recursion formulas for such a holomorphic solution yield

$$3\,!\,\gamma_{13} = \sum_{\nu=1}^{\infty} | a_{1\nu} | \sum_{\mu=1}^{\infty} | a_{\nu\mu} | \; | a_\mu |$$

and accordingly $\gamma_{13} = + \infty$ (cf. *loc. cit.*[2], II).

From Existence Theorem I and the above consistency theorems there follows, since our concepts are the non-linear generalizations of the concepts of Hilbert, the following:

Existence Theorem II.[24] *If* (31) *is bounded in* σ_r *the implicit system*

$$z_i = t\, \Phi_i (z_1, z_2, \cdots) ; \qquad (i = 1, 2, \cdots) \tag{54}$$

possesses in the circle $| t | < \alpha$, *where* α *is a suitably chosen constant, a holomorphic solution*

$$z_i = z_i(t) \tag{55}$$

and this solution (55) *lies in* σ_r *if* t *lies in the circle* $| t | < \alpha$.

The assumption that the Φ_i are independent of t, is no restriction on the generality of system (54) as we can adjoin to (54) the equation $z_0 = t\Phi_0$ where $\Phi_0 \equiv 1$.

If one chooses in (54) the Φ_i linear functions of $z_1, z_2, \cdots$ namely

$$\Phi_i = a_i + \sum_{j=1}^{\infty} a_{ij} z_j ; \qquad (i = 1, 2, \cdots) \tag{56}$$

then the vector (31) is then and only then bounded if the matrix $\| a_{ij} \|$ and the linear form (21) are both bounded in the sense of Hilbert and (54) becomes

$$z_i - t \sum_{j=1}^{\infty} a_{ij} z_j = t a_i ; \qquad (i = 1, 2, \cdots) \tag{54'}$$

so that Existence Theorem II becomes the Theorem of Hilb[25]: If the matrix $\| a_{ij} \|$ is bounded and if $\| \delta_{ij} \|$ denotes the unit matrix the matrix $\| \delta_{ij} - t a_{ij} \|$ has for sufficiently small values of $| t |$ a bounded reciprocal matrix which is a holomorphic function of t (Neumann's series of iterations).

The method by which the proof of the existence theorems proceeds is of such character that the generalization to the case where the components of (31) contain also parameters μ_i is almost trivial. (In this respect compare *loc. cit.*[2], I).

With the use of this parametric generalization there follows from Existence Theorem II by means of a device [26] the following theorem on the regular transformations of the Hilbert space into itself:

EXISTENCE THEOREM III.[27] *If*

$$\psi_1(z_1', z_2', \cdots), \quad \psi_2(z_1', z_2', \cdots), \cdots \tag{31'}$$

is a bounded vector in the domain

$$\sigma'_{r'}: \quad \sum_{k=1}^{\infty} |z_k'|^2 < r'^2$$

and if ψ_i contains no constant and no linear term there exists a sufficiently small r and a bounded vector (31) *of regular power series in σ_r such that the system*

$$z_i = z_i' + \psi_i(z_1', z_2', \cdots); \qquad (i = 1, 2, \cdots) \tag{57}$$

possesses in σ_r the inverse transformation

$$z_i' = \Phi_i(z_1, z_2, \cdots); \qquad (i = 1, 2, \cdots). \tag{58}$$

It follows from the Consistency Theorem IV that the restriction made in Existence Theorem III that the transformation of the Hilbert space into itself is to higher terms the identity transformation, is not an essential one, and Existence Theorem III is accordingly still valid if (57) be replaced by the more general system

$$z_i = \sum_{j=1}^{\infty} a_{ij} z_j' + \psi_i(z_1', z_2', \cdots); \qquad (i = 1, 2, \cdots) \tag{57'}$$

where $\| a_{ij} \|$ is a constant matrix (becoming in the normalized case (57) the unity matrix) which is bounded in the sense of Hilbert and which possesses a bounded reciprocal matrix, i. e. the linear approximation

$$z_i = \sum_{j=1}^{\infty} a_{ij} z_j' \tag{57a'}$$

of (57′) is a unique reversible transformation of the complex space of Hilbert into itself. In the case of finite many variables this assumption is obviously equivalent to the restriction that the Jacobian determinant be different from zero in the neighborhood of the origin.

In addition Existence Theorem II permits a corresponding generalization, i. e. Existence Theorem II remains true if we replace (54) by

$$(54') \qquad \sum_{j=1}^{\infty} a_{ij}z_j = t\,\Phi_i(z_1, z_2, \cdots); \qquad (i = 1, 2, \cdots)$$

provided that we assume the matrix $\| a_{ij} \|$ is bounded in the sense of Hilbert and possesses a bounded reciprocal $\| b_{ij} \|$.

Finally it should be mentioned that Existence Theorem III remains valid if the ψ_i depend not only upon the z_i' but also upon the z_i. The proof proceeds in exactly the same manner as for Existence Theorem III; as is also true for the generalization mentioned in connection with (57a').

THE JOHNS HOPKINS UNIVERSITY.

BIBLIOGRAPHY.

[1] The bibliography of this literature is to be found in the recently appearing book of F. R. Moulton, *Differential Equations*, New York, 1930, p. 375.

[2] In what follows, reference is made to the following papers: I. "Zur Theorie der unendlichen Differentialsysteme," *Mathematische Annalen*, Bd. 95 (1925), pp. 544-556; II. "Zur Lösung von Differentialsystemen mit unendlichvielen Verändlichen," *Mathematische Annalen*, Bd. 98 (1928), pp. 273-280; III. "Zur Analyse im Hilbertschen Raume," *Mathematische Zeitschrift*, Bd. 28 (1928), pp. 457-470. References to papers, in which applications are made of the existence theorems mentioned in the text, are given in the above papers and in the paper of Martin referred to in footnote (3).

[3] M. Martin, "Upon the Existence and Non-Existence of Isoenergetic Periodic Perturbations of the Undisturbed Circular Motions in the Restricted Problem of Three Bodies," *American Journal of Mathematics*, Vol. 34 (1931), pp. 259-273.

[4] Cf. pp. 1471-1476 of the article by E. Hellinger and O. Toeplitz in the *Encyklopädie der Mathematischen Wissenschaften*, Bd. 2III$_2$.

[5] Cf. E. Helly, "Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten," *Monatshefte für Mathematik und Physik*, Bd. 31 (1921), pp. 60-91.

[6] For the definitions of and the theory of convergence of the bounded linear forms (vectors) and bounded bilinear forms (matrices) of Hilbert, cf. E. Hellinger and O. Toeplitz, "Grundlagen für eine Theorie der unendlichen Matrizen," *Mathematische Annalen*, Bd. 69 (1910), pp. 289-330.

[7] This definition is not identical with the one incidentally proposed by Hilbert, "Wesen und Ziele einer Analysis der unendlichvielen unabhängigen Variablen," *Rendiconti del Circolo Matematico di Palermo*, Tomo 27 (1909), pp. 59-74, in which Hilbert considers the general power series not in the Hilbert space but in a more easily treated manifold.

[8] *Loc. cit.* (6), § 9, § 10.

[9] Cf. *loc. cit.* (2) III, pp. 454-455.

[10] Cf. *loc. cit.* (2) III, p. 457.

[11] Cf. *loc. cit.* (2) III, p. 455.

[12] Cf. *loc. cit.* (2) III, p. 456.

[13] The restriction (28), (28′) is not a sufficient condition for the boundedness of the matrix $\| a_{ij} \|$ since for example the condition is fulfilled for every skew-symmetric matrix and there exist skew-symmetric matrices which are not bounded.

One obtains a necessary and sufficient condition if one replaces (28) by [6]

$$\left| \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\zeta_i\bar{\zeta}_j \right| \leqq L.$$

[14] O. Toeplitz, "Über eine bei den Dirichletschen Reihen auftretende Aufgabe etc.," *Göttinger Nachrichten* (1913), pp. 417-432.

[15] Cf. *loc. cit.* ([9]).

[16] *Loc. cit.* ([2]) III, p. 464.

[17] W. L. Hart, "Linear Differential Equations in Infinitely Many Variables," *American Journal of Mathematics,* Vol. 39 (1917), pp. 407-424.

[18] For a more detailed discussion cf. *loc. cit* ([2]) II, p. 275-276; cf. also *loc. cit.* ([2]) III, p. 461, Fussnote [24].

[19] This domain is, in the case of finite differential systems, the one found by L. Fuchs.

[20] Cf. *loc. cit.* ([2]) I, p. 553-556.

[21] For a review on this subject cf. M. Müller, "Neuere Untersuchungen über den Fundamentalsatz in der Theorie der gewöhnlichen Differentialgleichungen," *Jahresbericht der Deutschen Mathematiker-Vereinigung,* Bd. 37 (1928), pp. 33-48.

[22] C. Carathéodory, *Vorlesungen über reelle Funktionen* (1918), pp. 665-668.

[23] In order to apply these methods it is necessary to define continuity by (16), (17) and not by (17), (18).

[24] *Loc. cit.* ([2]) III, p. 462.

[25] Cf. for example *loc. cit.* ([4]), p. 1431.

[26] Cf. *loc. cit.* ([2]) III, p. 468 or *Mathematische Annalen,* Bd. 96 (1926), p. 292 (Fussnote).

[27] Cf. a later appearing paper by Martin.

NOTE ON THE NUMERICAL VALUE OF A PARTICULAR MASS RATIO IN THE RESTRICTED PROBLEM OF THREE BODIES.

By Jenny E. Rosenthal.*

In order to fix the positions of the collinear libration points (L_1, L_2 and L_3 in the notation of E. Strömgren †) on the line joining the masses μ and $1-\mu$ we denote ρ_1 and $\bar{\rho}_1$ the distance of L_1 from the mass μ and the mass $1-\mu$ respectively, by ρ_2 the distance of L_2 from the mass μ, and by ρ_3 the

distance of L_3 from the mass $1-\mu$. In a recent paper Martin ‡ has shown that the distances ρ_1, $\bar{\rho}_1$, ρ_2 and ρ_3 are monotone functions of μ in the interval $0 < \mu < 1$ and that they possess the following property: In this interval there exists one and only one value μ^* of μ for which

$$(1) \qquad \rho_1(\mu) \lesseqqgtr \rho^* \lesseqqgtr \rho_2(\mu) \quad \text{according as} \quad \mu \lesseqqgtr \mu^*.$$

It follows from considerations of symmetry that

$$(2) \qquad \bar{\rho}_1(\mu) \lesseqqgtr \rho^* \lesseqqgtr \rho_3(\mu) \quad \text{according as} \quad \mu \gtreqqless 1-\mu^*.$$

The equations for the determination of ρ^* and μ^* have been derived by Martin ‡ from which the author has calculated the numerical values of ρ^* and μ^* finding

$$\rho^* = 0.938933, \qquad \mu^* = 0.9992718, \qquad 1-\mu^* = 0.0007282;$$

where the results are accurate to the last significant figure indicated.

For the mass ratio of the Sun and Jupiter we have approximately $\mu = 0.9990 < \mu^*$. Since the libration points L_1, L_2 and L_3 are separated by the masses it follows from the inequalities (1) that the libration point nearest the Sun is to be found between the Sun and Jupiter.

The Johns Hopkins University.

---

* National Research Fellow.

† E. Strömgren, *Publikationer og mindre Meddelelser fra Köbenhavns Observatorium*, Nr. 39 (1922).

‡ M. Martin, *American Journal of Mathematics*, Vol. 53 (1931), pp. 167-179.

UPON THE EXISTENCE AND NON-EXISTENCE OF ISOENERGETIC PERIODIC PERTURBATIONS OF THE UNDISTURBED CIRCULAR MOTIONS IN THE RESTRICTED PROBLEM OF THREE BODIES.

By Monroe Martin.

Introduction. In the restricted problem of three bodies if we denote the masses of the two bodies, rotating in concentric circles about their center of mass, by μ and $1-\mu$ then for $\mu=0$ a possible orbit for the third body is a circular motion about the non-zero mass. In this paper we treat the following problem: Does there exist, for sufficiently small, positive values of μ, and for a given value of the constant of relative energy (Jacobian constant), an isoenergetic † series of periodic solutions of the equations of motion, the members of which converge to the circular motion mentioned above, and the periods of which converge to the period of this circular motion, when the parameter μ of the series converges to zero? Denoting by n the angular velocity, in the non-rotating coördinate sytem, of the third body in its circular motion for $\mu=0$, Levi-Civita,‡ and Birkhoff,§ employing the methods of Poincaré,¶ have shown that, when n is not the ratio of two successive integers, such an isoenergetic series of periodic orbits actually exists. For n the ratio of two successive integers the problem is a resonance problem and neither the existence or non-existence of this isoenergetic series of periodic orbits has as yet been mathematically demonstrated.‖ The existence of this isoenergetic series for these critical values of n requires the vanishing of

† Isoenergetic is here taken to mean that the Jacobian constant is independent of μ. For the corresponding isoperiodic case in which the Jacobian constant is a function of μ, cf. A. Wintner, "Über eine Revision der Sortentheorie des restringierten Dreikörperproblems," *Sitzungsberichte der Sächsischen Akademie der Wissenschaften zu Leipzig*, Vol. 82 (1930), pp. 3-56.

‡ T. Levi-Civita, "Sopra alcuni criteri di instabilità," *Annali di Matematica* (3), Vol. 5 (1901), pp. 282-289.

§ G. D. Birkhoff, "The Restricted Problem of Three Bodies," *Rendiconti del Circolo Matematico di Palermo*, Vol. 39 (1915), § 11; and also, "Dynamical Systems," *American Mathematical Society Colloquium Publications*, Vol. 9, New York (1927), pp. 139-143.

¶ H. Poincaré, *Méthodes Nouvelles de la Mécanique Céleste*, Vol. 1 (1892), pp. 79-119.

‖ Concerning a paper of Poincaré in the *Bulletin Astronomique*, cf. § 6, A. Wintner, *loc. cit.*

certain expressions, the vanishing of which expresses the fulfillment of a type or orthogonality condition (Verzweigungsgleichungen).† It will be demonstrated that for sufficiently small values of μ this condition is not fulfilled, and consequently the isoenergetic series of periodic orbits cannot exist for these critical values of n. The mathematical apparatus employed differs from that employed by Levi-Civita and Birkhoff, the method here being that of the infinitely many variables,‡ as developed for the problems of celestial mechanics in the papers of Wintner.§ In order to put in evidence the method used in the proof of the non-existence of an isoenergetic series for the critical values of n, it is necessary to present the existence proof, with the method of infinitely many variables, for the non-critical values of n.

1. *The non-linear differential equation of the normal perturbation.* If the origin of the rotating system of coördinates be taken at the mass $1-\mu$, the Lagrangian function for the third mass may be written

$$L \equiv \tfrac{1}{2}(\dot{x}^2+\dot{y}^2)+(x\dot{y}-\dot{x}y)+F(x,y;\mu), \tag{1}$$

where we have placed

$$F(x,y;\mu) \equiv \tfrac{1}{2}[(x-\mu)^2+y^2] + \mu[(x-1)^2+y^2]^{-1/2}+(1-\mu)(x^2+y^2)^{-1/2}.$$

The dot is used here and throughout the paper to denote differentiation with respect to the time. In order to obtain the differential equation of the normal displacement we introduce polar coördinates ρ and τ and it will be convenient to place $\rho = a+\zeta$, accordingly

$$x=(a+\zeta)\cos\tau, \qquad y=(a+\zeta)\sin\tau. \tag{2}$$

The Lagrangian function and F when expressed in terms of the new

† Cf. the papers referred to in the footnotes § and ‡ below.

‡ For another method developed by Lichstenstein in order to treat non-linear boundary value problems with the use of Greenian matrices and successive approximations and which would be in the present case equivalent, cf. L. Lichstenstein, "Zur Maxwellschen Theorie der Saturnringe," *Mathematische Zeitschrift* 17 (1923), pp. 62-110, and the further developments of this method by E. Hölder, "Mathematische Untersuchungen zur Himmelsmechanik," *Mathematische Zeitschrift*, Vol. 31 (1930), pp. 225-239.

§ Cf. for instance A. Wintner, "Über die Differentialgleichungen der Himmelsmechanik," *Mathematische Annalen*, 96 (1926), pp. 284-312.

coördinates ζ and τ will be denoted by L^* and F^* respectively. On expressing (1) in terms of ζ and τ by means of (2) we obtain

$$(3)\quad \begin{aligned} L^* &\equiv \tfrac{1}{2}[\dot{\zeta}^2 + (a+\zeta)^2\dot{\tau}^2] + (a+\zeta)^2\dot{\tau} + F^*(\zeta, \tau; \mu, a), \\ F^*(\zeta, \tau; \mu, a) &\equiv \tfrac{1}{2}(a+\zeta)^2 + (a+\zeta)^{-1} + \mu^2/2 \\ &+ \mu\{[1 - 2a\cos\tau + a^2 + 2(a - \cos\tau)\zeta + \zeta^2]^{-\frac{1}{2}} - (a+\zeta)\cos\tau - (a+\zeta)^{-1}\} \end{aligned}$$

from which the Lagrangian equations

$$(4)\qquad \frac{d}{dt}\frac{\partial L^*}{\partial \dot{\zeta}} - \frac{\partial L^*}{\partial \zeta} = 0, \quad \frac{d}{dt}\frac{\partial L^*}{\partial \dot{\tau}} - \frac{\partial L^*}{\partial \tau} = 0,$$

which possess the Jacobian integral

$$(5)\qquad \dot{\zeta}^2 + (a+\zeta)^2\dot{\tau}^2 = 2F^* - C,$$

may be derived.

For $\mu = 0$ equations (4) take the following simple form:

$$(6)\quad \ddot{\zeta} - (1+\dot{\tau})^2(a+\zeta) + (a+\zeta)^{-2} = 0, \qquad \frac{d}{dt}[(a+\zeta)^2(1+\dot{\tau})] = 0,$$

and possess the solution

$$(7)\qquad \zeta \equiv 0, \qquad \tau = (n-1)t + \tau_0,$$

in which n (the angular velocity of the third mass in the non-rotating coördinate system) is determined from a, the radius of the circular orbit, by the third law of Kepler

$$(8)\qquad n^2a^3 = 1.$$

Putting $\mu = 0$ in the Jacobian integral (5) and inserting the solution (7), the Jacobian constant C, for the circular orbits (7) is obtained in terms of a, the radius of the circular orbit, as follows:

$$(9)\qquad C = 2a^{\frac{1}{2}} + 1/a.$$

For a given value of the Jacobian constant C we may employ (5) to replace (3) by a new Lagrangian function Λ, homogeneous of degree one in $\dot{\zeta}$ and $\dot{\tau}$, namely

$$(10)\qquad \Lambda \equiv \{[2F^* - C][\dot{\zeta}^2 + (a+\zeta)^2\dot{\tau}^2]\}^{\frac{1}{2}} + (a+\zeta)^2\dot{\tau}.$$

Since Λ is homogeneous of degree one in $\dot{\zeta}$ and $\dot{\tau}$ the Lagrangian equations

$$(11)\qquad \frac{d}{dt}\frac{\partial \Lambda}{\partial \dot{\zeta}} - \frac{\partial \Lambda}{\partial \zeta} = 0, \quad \frac{d}{dt}\frac{\partial \Lambda}{\partial \dot{\tau}} - \frac{\partial \Lambda}{\partial \tau} = 0,$$

are invariant under a change of independent variable

$$(12) \qquad t = t(\theta), \qquad d\tau/d\theta > 0.$$

Having decided upon a suitable choice for the independent variable θ the principle of Maupertuis † states that if we write the solutions of (11) with θ as independent variable in the form

$$(13) \qquad \zeta = \zeta(\theta), \qquad \tau = \tau(\theta),$$

then every solution of (4) for the given value of the Jacobian constant mentioned above, can be written in the form

$$(14) \qquad \zeta = \zeta(\theta(t)), \qquad \tau = \tau(\theta(t)).$$

The function $\theta(t)$ is the inverse of the function (12) and is obtained by a quadrature from

$$(15) \qquad \left(\frac{d\theta}{dt}\right)^2 = \frac{2F^*(\zeta, \tau; \mu, a) - C}{\left(\dfrac{d\zeta}{d\theta}\right)^2 + (a+\zeta)^2\left(\dfrac{d\tau}{d\theta}\right)^2}.$$

Since θ is arbitrary we may take $\theta = \tau$, that is we now regard τ, originally a dependent variable, as the independent variable and seek to express the variable ζ as a function of it. The Lagrangian function, with τ the independent variable, will be denoted by Λ^* and we have from (10)

$$(16) \qquad \Lambda^* = \{[2F^* - C][\zeta'^2 + (a+\zeta)^2]\}^{1/2} + (a+\zeta)^2; \qquad ' = d/d\tau.$$

The Lagrangian equation, the solution of which yields ζ as a function of τ and which will be called the equation of the normal perturbation is accordingly

$$(17) \qquad \frac{d}{d\tau}\frac{\partial\Lambda}{\partial\zeta'} - \frac{\partial\Lambda}{\partial\zeta} = 0.$$

Equation (15) takes the form

$$(18) \qquad \frac{d\tau}{dt} = \phi(\tau)$$

and serves to determine τ as a function of t.

It is clear, since the principle of Maupertius proceeds on the assumption that the Jacobian constant has a given value, that in order to obtain an isoenergetic series (with μ as parameter) of solutions of (4), the members of which converge to the circular motion (7) as μ converges to zero, it is not only sufficient but also necessary to obtain the solutions of (17) and (18).

† G. D. Birkhoff, "Dynamical Systems with Two Degrees of Freedom," *Transactions of the American Mathematical Society*, Vol. 18 (1917), pp. 202-204 or F. D. Murnaghan, "The Principle of Maupertuis," *Proceedings of the National Academy of Sciences*, Vol. 17 (1931), pp. 128-132.

If one is not interested in the history of the third mass in its motion, but only the form of its orbit, equation (18) may be discarded.

2. *The calculation of the equation of the normal perturbation.* We shall now treat the equation of the normal perturbation in some detail. If we substitute the expression for Λ^* given by (16) in equation (17) we obtain

$$(19) \qquad \zeta'' - \frac{2[(a+\zeta)^2+\zeta'^2]^{3/2}}{(a+\zeta)(2F^*-C)^{1/2}}$$
$$+ \frac{F_\tau^*[(a+\zeta)^2+\zeta'^2]\zeta' - F_\zeta^*[(a+\zeta)^2+\zeta'^2][a+\zeta]^2}{(2F^*-C)(a+\zeta)^2} - \frac{2\zeta'^2}{a+\zeta} = 0,$$

or

$$(19') \qquad \zeta'' = \Phi(\mu, \zeta, \zeta', \cos\tau; \mu, a),$$

where we understand $F_\tau^* = \partial F^*/\partial\tau$ and $F_\zeta^* = \partial F^*/\partial\zeta$.

We treat, once and for all, only those values of a for which

$$(20) \qquad 0 < a < \infty, \quad a \neq 1.$$

We have then

$$(21) \qquad 1 - 2a\cos\tau + a^2 \geqq (1-a)^2 > 0; \qquad -\infty < \tau < \infty,$$

and from (9) and (3)

$$(22) \qquad 2F^*(0,\tau;0,a) - C \equiv a^2(a^{-3/2}-1)^2 > 0; \qquad -\infty < \tau < +\infty.$$

We wish to emphasize here, that while Φ is a function of a, the value of a is fixed, inasmuch as the Jacobian constant C is assumed to have been given [cf. (9)]. Since a is fixed it follows from (19′), (19), (20), (21), (22) that the function Φ is regular for sufficiently small values of $|\zeta|$, $|\zeta'|$ and $|\mu|$ and for all real values of τ. Consequently there exists two positive numbers R and K such that in the domain

$$(23) \qquad |\zeta| < R, \quad |\zeta'| < R, \quad |\mu| < R, \quad -\infty < \tau < +\infty,$$

we have

$$(24) \qquad \left|\frac{\partial^2\Phi}{\partial\tau^2}\right| < K$$

and the convergent developments

$$(25) \qquad \Phi = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} A_{jkl}\,\mu^j\zeta^k\zeta'^l,$$

$$(26) \qquad \frac{\partial^2\Phi}{\partial\tau^2} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{d^2A_{jkl}}{d\tau^2}\,\mu^j\zeta^k\zeta'^l,$$

where from (19) the coefficients A_{jkl} of (25) are obviously real and even functions of τ and can be expanded in a regular Fourier series of the form

(27) $$A_{jkl} = \sum_{\nu=-\infty}^{\infty} A_{jkl}^{(\nu)} e^{\nu\tau i}, \qquad A_{jkl}^{(\nu)} = A_{jkl}^{(-\nu)} \gtreqless 0; \qquad (j, k, l = 0, 1, 2, \cdots).$$

The inequality (24) for the function (26) being valid it follows by Cauchy's theorem on convergent power series that in the domain (23) we have

(28) $$\left| \frac{d^2 A_{jkl}}{d\tau^2} \right| < KR^{-(j+k+l)}; \qquad (j, k, l = 0, 1, 2, \cdots).$$

Since (27) is a regular Fourier series, it follows on differentiation term by term

(29) $$\frac{d^2 A_{jkl}}{d\tau^2} = - \sum_{\nu=-\infty}^{\infty} \nu^2 A_{jkl}^{(\nu)} e^{\nu\tau i}; \qquad (j, k, l = 0, 1, 2, \cdots).$$

Since the absolute values of the Fourier coefficients in the development of a periodic function are not greater than the maximum of the absolute value of the function we have from (28) and (29) †

(30) $$| A_{jkl}^{(\nu)} | < \nu^{-2} KR^{-(j+k+l)}; \qquad (j, k, l = 0, 1, 2, \cdots; \nu = 0, \pm 1, \pm 2, \cdots).$$

We now proceed to calculate explicitly the following coefficients in the expansion (25):

(31) $$A_{jkl} \quad \text{for} \quad j, k, l \leqq 1.$$

The labor of this calculation is lessened if instead of proceeding directly from (19) we develope the Lagrangian function Λ^* according to powers of μ, ζ and ζ' and obtain the coefficients of the terms ζ, ζ^2, ζ'^2 and $\mu\zeta$ (since Λ^* is a function of ζ'^2 it is clear at once that the coefficients of ζ', $\mu\zeta'$ and $\zeta\zeta'$ are zero). On taking the Lagrangian derivative of the development of Λ^* the coefficients (31) may be obtained immediately. From (16) and (3) we obtain

(32) $$\Lambda^* = [(A + B\mu + \mu^2)(a^2 + 2a\zeta + \zeta^2 + \zeta'^2)]^{1/2} + (a + \zeta)^2,$$

where we have placed

(33) $$A = (a + \zeta)^2 + 2(a + \zeta)^{-1} - C,$$

(34) $$B = 2[r^2 + 2(a - \cos\tau)\zeta + \zeta^2]^{-1/2} - 2(a + \zeta)\cos\tau - 2(a + \zeta)^{-1},$$

in which r denotes the distance between the mass μ and the third mass, i. e.

(35) $$r^2 = 1 - 2a\cos\tau + a^2.$$

If we develop A and B (for sufficiently small $|\zeta|$) according to powers of ζ, and replace C by (9), we obtain for the first terms mentioned above in the development of A and B respectively:

---

† If we agree to write $1/0 = 1$ the validity of the inequalities (30) for $\nu = 0$ follows directly from (25).

$$(36) \qquad A_2 = (a^{-\frac{1}{2}} - a)^2 + 2(a - a^{-2})\zeta + (1 + 2a^{-3})\zeta^2$$

$$(37) \qquad B_1 = 2(r^{-1} - a^{-1} - a\cos\tau) + 2[a^{-2} + (\cos\tau - a)r^{-3} - \cos\tau]\zeta.\dagger$$

As a first step in the development of (32) according to powers of μ, ζ and ζ' we have (for sufficiently small $|\zeta|$ and $|\zeta'|$)

$$(38) \qquad \Lambda^* = a(A + B\mu + \mu^2)^{\frac{1}{2}}(1 + a^{-1}\zeta + a^{-2}\zeta'^2/2) + (a+\zeta)^2 + \Lambda_1^*,$$

where Λ_1^* contains terms of order higher than the second in ζ and ζ'. On developing (for sufficiently small $|\mu|$) according to powers of μ we have on replacing A by A_2 and B by B_1

$$(39) \qquad \Lambda^* = aA_2^{\frac{1}{2}}(1 + a^{-1}\zeta + \tfrac{1}{2}a^{-2}\zeta'^2) \\ + (a+\zeta)^2 + \tfrac{1}{2}aA_2^{-\frac{1}{2}}B_1(1 + a^{-1}\zeta + \tfrac{1}{2}a^{-2}\zeta'^2)\mu + \Lambda_2^*,$$

where Λ_2^* contains terms of order higher than the second in μ, ζ and ζ'. From (36) we have

$$(40) \qquad A_2^{\frac{1}{2}} = \boldsymbol{A}\{1 + a^{-2}\boldsymbol{A}^{-2}(a^3-1)\zeta \\ + \tfrac{1}{2}[(a^3+2)a\boldsymbol{A}^2 - (a^3-1)^2]a^{-4}\boldsymbol{A}^{-4}\zeta^2\} + \cdots,$$

where we have written

$$(41) \qquad \boldsymbol{A} = a^{-\frac{1}{2}} - a.$$

From (37) and (40)

$$(42) \qquad \tfrac{1}{2}aA_2^{-\frac{1}{2}}B_1 = \boldsymbol{A}^{-1}\{ar^{-1} - a^2\cos\tau - 1 + [a^{-1} + ar^{-3}(\cos\tau - a) \\ - a\cos\tau - (ar^{-1} - a^2\cos\tau - 1)(a^3-1)a^{-2}\boldsymbol{A}^{-2}]\zeta\} + \cdots.$$

Equation (39) may now be written

$$(43) \qquad \Lambda^* = \tfrac{1}{2}a^{-1}\boldsymbol{A}\zeta'^2 + \tfrac{1}{2}[2a^3\boldsymbol{A}^3 + 3a^4\boldsymbol{A}^2 - (a^3-1)^2]a^{-3}\boldsymbol{A}^{-3}\zeta^2 \\ + [a^3 + a\boldsymbol{A}^2 + 2a^2\boldsymbol{A} - 1]a^{-1}\boldsymbol{A}^{-1}\zeta \\ + [(a\boldsymbol{A}^2 - a^3 + 1)ar^{-1} + (\cos\tau - a)a^3\boldsymbol{A}^2r^{-3} \\ + a^2(a^3 - 2a\boldsymbol{A}^2 - 1)\cos\tau + a^3 - 1]a^{-2}\boldsymbol{A}^{-3}\mu\zeta + \Lambda_3^*,$$

where Λ_3^* contains terms of higher order than the second in μ, ζ and ζ', together with unessential terms independent of ζ and ζ'. On multiplying (43) throughout by $a\boldsymbol{A}^{-1}$ we obtain

$$(44) \qquad a\boldsymbol{A}^{-1}\Lambda^* = \Lambda^*_0 + a\boldsymbol{A}^{-1}\Lambda^*_3,$$

where from (8) and (41)

$$(45) \qquad \Lambda^*_0 \equiv \tfrac{1}{2}\zeta'^2 - \tfrac{1}{2}n^2(n-1)^{-2}\zeta^2 + \mu\zeta A_{100}$$

† Since B_1 occurs in the expansion of Λ^* multiplied by μ [cf. (39) following] it is sufficient to give here only the linear and constant terms in the development of B_1 according to powers of ζ.

in which we have placed

$$(46)\quad A_{100} \equiv [(aA^2 - a^3 + 1)ar^{-1} + (\cos\tau - a)a^3A^2r^{-3} + a^2(a^3 - 2aA^2 - 1)\cos\tau + a^3 - 1]a^{-1}A^{-4}.$$

We now introduce the parameter of Hill

$$(47)\qquad m = (n-1)^{-1},$$

and note at this point that a, n and m by virtue of (8), (9) and (47) are all determined by the Jacobian constant C. With the exception of $m = 0$, -1 and $\pm\infty$, which have been excluded by (20), the integral values of m, namely,

$$(48)\qquad m = 1, \pm 2, \pm 3, \pm 4, \cdots,$$

correspond to the critical values of n mentioned in the introduction, and conversely. Expressing a and n in terms of m equations (45) and (46) become respectively:

$$(49)\qquad \Lambda^*_0 \equiv \tfrac{1}{2}\zeta'^2 - \tfrac{1}{2}(m+1)^2\zeta^2 + A_{100}\,\mu\zeta,$$

$$(50)\quad A_{100} \equiv m^{4/3}(m+1)^{2/3}\{2(m+1)r^{-1} - m^{2/3}(m+1)^{-2/3}(2m+3)\cos\tau + m^{2/3}(m+1)^{-2/3}r^{-3}\cos\tau - m^{4/3}(m+1)^{-4/3}r^{-3} - m^{-2/3}(m+1)^{2/3}(2m+1)\}.$$

The equation of normal perturbation (19) may now be written

$$(51)\qquad \zeta'' + (m+1)^2\zeta = \mu A_{100} + \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} A_{jkl}\,\mu^j\zeta^k\zeta'^l,$$

or from (27)

$$(51')\quad \zeta'' + (m+1)^2\zeta = \mu \sum_{\nu=-\infty}^{\infty} A_{100}^{(\nu)} e^{\nu\tau i} + \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{\nu=-\infty}^{\infty} A_{jkl}^{(\nu)} e^{\nu\tau i}\mu^j\zeta^k\zeta'^l,$$

where in (51) and (51)$'$

$$(52)\qquad j + k + l \geqq 2.$$

3. *The infinite system of conditions for the existence of a periodic solution of the equation of the normal perturbation.* We have seen in § 1 that in order to obtain an isoenergetic series (with μ as parameter) of solutions of the differential equations (4), the members of which converge to the circular motion (7) as μ converges to zero, it is both necessary and sufficient to consider the solutions of (17). If we are to treat the series of periodic solutions of (4) mentioned in the introduction it is necessary in addition, as we see from (2), that ζ be a periodic function of period 2π in τ, that is the period of ζ with respect to τ is independent of μ. On the other hand

the period of ζ with respect to t is determined from (18) and will obviously depend upon μ. Such a solution for ζ, real and periodic with period 2π in τ, will be sought for in the form of the Fourier development

$$\zeta(\tau)=\sum_{h=-\infty}^{\infty} z_h e^{h\tau i},$$

where it will be convenient to introduce the substitutions †

$$(53)\qquad z_h=x^3h^{-4}y_h,\qquad \mu=x^4;\qquad\qquad (h=0\pm 1,\pm 2,\cdots),$$

so that we have

$$(54)\qquad \zeta(\tau)=x^3\sum_{h=-\infty}^{\infty} h^{-4}y_h e^{h\tau i}.$$

The infinitely many unknowns y_k are, for a fixed value of the Jacobian constant C (and hence of a or m), functions of x which are to be so determined that for sufficiently small values of x the series (54) converges and is a solution of the differential equation (51), becoming identically zero for $x=0$.

We avail ourselves of the notation

$$(55)\qquad [\sum_{\nu=-\infty}^{\infty} a_\nu e^{\nu\tau i}]_j=a_j;\ddagger\qquad\qquad (j=0,\pm 1,\pm 2,\cdots)$$

and the infinite system of conditions which the y_k must satisfy in order that (54) be a solution of (51) may be made to take the form

$$(56)\quad \{h^{-2}(m+1)^2-1\}y_h=x\,\psi_h(x;y_0,y_1,y_{-1},\cdots);\quad (h=0,\pm 1,\pm 2,\cdots),$$

where we have placed in the notation (55)

$$(57)\qquad \psi_h=\sum_{j=0}^{\infty}{}'\sum_{k=0}^{\infty}{}'\sum_{l=0}^{\infty}{}' h^2x^{4j-4}[A_{jkl}\zeta^k\zeta'^l]_h;\S\qquad (h=0,\pm 1,\pm 2,\cdots)$$

and which can be expanded as follows:

$$(57')\qquad \psi_h=\sum_{j=0}^{\infty}{}'\sum_{k=0}^{\infty}{}'\sum_{l=0}^{\infty}{}'\sum_{p=-\infty}^{\infty}\sum_{q=-\infty}^{\infty} h^2x^{4j-4}A_{jkl}^{(h-p)}\,[\zeta^k]_{p-q}[\zeta'^l]_q;$$
$$(h=0,\pm 1,\pm 2,\cdots).$$

The ψ_h are power series in the infinitely variables and we now proceed

† Here and throughout the remainder of the paper we shall retain the convention agreed upon in the footnote on page 264.

‡ If $f_1, f_2,\cdots,f_n$ be any n Fourier series we shall understand by the symbol $[f_i\cdot f_j]_k$ the coefficient of the k-th term in the Fourier series obtained by the multiplication of the two Fourier series f_i and f_j. If $c_1, c_2,\cdots c_n$ be any n constants the symbol $[\ ]_k$ possesses the property $[c_1f_1\cdot c_2f_2\cdots c_nf_n]_k=c_1c_2\cdots c_n[f_1\cdot f_2\cdots f_n]_k$.

§ The prime affixed to the summation sign indicates that the summation labels take the following values: $j=1,\ k=0,\ l=0;\ j+k+l\geqq 2$.

to establish an inequality for the best majorant † $\bar{\psi}_k$. Before doing so we state here a few facts on the multiplication of a special type of Fourier series.‡ In this connection we have the inequality §

$$(58)\qquad \sum_{k=-\infty}^{\infty} k^{-2}(l-k)^{-2} < 17l^{-2}; \qquad (l=0,\pm 1,\pm 2,\cdots).$$

The content of this inequality is that for n Fourier series f_j $(j=1,2,\cdots n)$ where in the notation (55)

$$(59)\qquad |[f_j]_k| \leqq k^{-2}; \qquad (j=1,2,\cdots n;\ k=0,\pm 1,\pm 2,\cdots),$$

the following inequalities exist

$$(60)\qquad |[f_i\cdot f_j]_k| < 17k^{-2}; \qquad (i,j=1,2,\cdots n,\ k=0,\pm 1,\pm 2,\cdots),$$

and in general if $C_1, C_2,\cdots C_n$ denote any n constants whatsoever

$$(61)\qquad |[C_1f_1\cdot C_2f_2\cdots C_nf_n]_k| < |C_1C_2\cdots C_n|\, 17^{n-1}k^{-2}; \qquad (k=0,\pm 1,\pm 2,\cdots)$$

because, for example

$$|[C_1f_1\cdot C_2f_2]_k| = |C_1C_2[f_1\cdot f_2]_k| < |C_1C_2|\, 17k^{-2} \quad (k=0,\pm 1,\pm 2,\cdots).$$

We now show that for the functions ψ_h defined by (57) there exists two positive numbers $\boldsymbol{a}$ and M_1 so that the best majorant $\bar{\psi}_h$ satisfies the inequality

$$(62)\qquad \bar{\psi}_h(\boldsymbol{a};1,1,\cdots) < M_1; \qquad (h=0,\pm 1,\pm 2,\cdots).$$

We introduce the notation

$$(63)\qquad \zeta^* = |x|^3 \sum_{j=-\infty}^{\infty} |j^{-4}y_j|\, e^{j\tau i}, \quad \zeta'^* = |x|^3 \sum_{j=-\infty}^{\infty} |j^{-3}y_j|\, e^{j\tau i}$$

$$A^*_{jkl} = \sum_{\nu=-\infty}^{\infty} |A^{(\nu)}_{jkl}|\, e^{\nu\tau i}, \qquad (j,k,l=0,1,2,\cdots).$$

It follows from (61) that in the domain

$$(64)\qquad |y_j| \leqq 1; \qquad (j=0,\pm 1,\pm 2,\cdots),$$

the following inequalities are valid:

$$(65)\qquad [\zeta^{*k}]_j < j^{-2}|x|^{3k}17^{k-1}, \quad [\zeta'^{*l}]_j < j^{-2}|x|^{3l}17^{l-1}; \qquad (k=0,1,2,\cdots;\ j=0,\pm 1,\pm 2,\cdots)$$

† By the best majorant of a function defined by a power series we shall understand the function defined by the power series obtained on replacing the coefficients of the original power series by their absolute values.

‡ A. Wintner, *loc. cit.*, footnote § p. 260.

§ A. Wintner, "Über die Konvergenzefragen der Mondtheorie," *Mathematische Zeitschrift*, Vol. 30 (1929), pp. 219-220.

and therefore from (30) and (63)

$$(66)\quad [A^*_{jkl}\,\zeta^{*k}\zeta'^{*l}]_h < \frac{K}{h^2}\;\frac{17^{k+l}\,|\,x\,|^{3(k+l)}}{R^{j+k+l}} \leqq \frac{K}{h^2}\left(\frac{17}{R}\right)^{j+k+l}|\,x\,|^{3(k+l)};$$
$$(j,k,l=0,1,2,\cdots;\ h=0,\pm1,\pm2,\cdots).$$

From (57) we have

$$(67)\quad \bar{\psi}_h(x;y_0,y_1,\cdots)\leqq\psi^*_h(x;y_0,y_1,\cdots)$$
$$\equiv\sum_{j=0}^{\infty}{}'\sum_{k=0}^{\infty}{}'\sum_{l=0}^{\infty}{}'\,h^2\,|\,x\,|^{4j-4}\,[A^*_{jkl}\,\zeta^{*k}\zeta'^{*l}]_h;\quad (h=0,\pm1,\pm2,\cdots),$$

and from (66)

$$(68)\quad \psi_h^*(x;1,1,\cdots)<\Psi(x)\equiv K\sum_{j=0}^{\infty}{}'\sum_{k=0}^{\infty}{}'\sum_{l=0}^{\infty}{}'\left(\frac{17\,|\,x\,|}{R}\right)^{j+k+l}|\,x\,|^{(3j+2k+2l-4)};$$
$$(h=0,\pm1,\pm2,\cdots).$$

It follows that $\Psi(x)$ is a regular power series in $|\,x\,|$. It follows by Cauchy's test that if

$$(69)\qquad |\,x\,|<\min\,(R/17,1)\dagger$$

and series $\Psi(x)$ is convergent. Consequently there exist two positive numbers $\boldsymbol{a}$ and M_1 so that (62) is valid.

4. *Existence proof for the non-critical case* $m\not\equiv 0\pmod 1$. Before entering upon the existence proof it is necessary to state an existence theorem on infinite systems.‡

THEOREM I. *Given the power series*

$$(70)\qquad f_{pq}(x_1,x_2,\cdots,x_s;\ y_1,y_2,\cdots;\ \lambda_1,\lambda_2,\cdots)$$
$$(p=1,2,\cdots,s;\ q=1,2,\cdots),$$

in infinitely many variables so that there exist positive numbers

$$(71)\qquad \alpha;\ \boldsymbol{a};\ b_1,b_2,\cdots;\ \Lambda_1,\Lambda_2,\cdots;\ M_1,M_2,\cdots,$$

for which, if $\bar{f}_{pq}$ *denote the best majorant* § *of* f_{pq}, *the following inequalities exist*

$$(72)\qquad \alpha<b_j/M_j;\ \bar{f}_{pj}\,(\boldsymbol{a},\boldsymbol{a},\cdots;\ b_1,b_2,\cdots;\ \Lambda_1,\Lambda_2,\cdots)\leqq M_j,$$
$$(p=1,2,\cdots s;\ j=1,2,\cdots),$$

then the system

† The symbol min (a,b) signifies the minimum of the two numbers a and b.
‡ A. Wintner, *loc. cit.* (8), pp. 291-294.
§ Cf. footnote †, p. 268

$$(73) \qquad y_j = \sum_{p=1}^{s} x_p f_{pj}; \qquad (j = 1, 2, \cdots),$$

possesses in the domain

$$(74) \qquad |x_p| < \min(\alpha/s, \boldsymbol{a}); \quad |\lambda_k| < \Lambda_k; \quad (p = 1, 2, \cdots s;\ k = 1, 2, \cdots),$$

one and only one solution $y_j(x_1, x_2, \cdots, x_s; \lambda_1, \lambda_2, \cdots)$. *This solution is holomorphic in the domain* (74) *and satisfies the inequalities*

$$(75) \qquad |y_i(x_1, x_2, \cdots, x_s; \lambda_1, \lambda_2, \cdots)| < b_i \qquad (i = 1, 2, \cdots),$$

and of course

$$(76) \qquad y_i(0, 0, \cdots 0; \lambda_1, \lambda_2, \cdots) \equiv 0 \qquad (i = 1, 2, \cdots).$$

If the power series (70) *and the arguments are real the solution* $y_j(x_1, x_2, \cdots x_s; \lambda_1, \lambda_2, \cdots)$ *is also real.*

If we place $s = 1$, $M_i = M$, $b_i = b$, $(i = 1, 2, \cdots)$ and reject the parameters λ we have the following special case of Theorem I:

THEOREM Ia. *Given the power series* $f_j(x; y_1, y_2, \cdots)$ *in infinitely many variables so that there exist three positive numbers*

$$(77) \qquad \boldsymbol{a},\ b,\ M,$$

for which, if $\bar{f}_j$ *denote the best majorante of* f_j, *the following inequalities exist*

$$(78) \qquad \bar{f}_j(\boldsymbol{a}; b, b, \cdots) < M \qquad (j = 1, 2, \cdots),$$

then the system

$$(79) \qquad y_i = x f_i; \qquad (i = 1, 2, \cdots),$$

possesses in the domain

$$(80) \qquad x < \min(b/M, \mathbf{a}),$$

one and only one solution $y_j(x)$. *This solution is holomorphic in the domain* (80) *and satisfies the inequalities*

$$(81) \qquad |y_j(x)| < b; \qquad (j = 1, 2, \cdots),$$

and of course

$$(82) \qquad y_j(0) = 0; \qquad (j = 1, 2, \cdots).$$

If the power series $f_j(x; y_1, y_2, \cdots)$ *and the arguments are real the solution* $y_j(x)$ *is also real.*

If $m \not\equiv 0 \pmod 1$ the infinite system (56) can be written

$$(83) \qquad y_j = x f_j(x; y_0, y_1, y_{-1}, \cdots); \qquad (j = 0, \pm 1, \pm 2, \cdots),$$

where we have placed

$$(84)\qquad f_j = [(m+1)^2/j^2 - 1]^{-1}\psi_j; \qquad (j = 0, \pm 1, \pm 2, \cdots),$$

and since $m \not\equiv 0 \pmod 1$ there exists a positive number M_2 so that

$$(85)\qquad |\,[(m+1)^2/j^2 - 1]^{-1}\,| < M_2; \qquad (j = 0, \pm 1, \pm 2, \cdots).$$

We have from (62), (84) and (85)

$$(86)\qquad \bar{f}_j(\boldsymbol{a};\, 1, 1, \cdots) < M; \qquad (j = 0, \pm 1, \pm 2, \cdots),$$

on putting $M = M_1 M_2$. From Theorem Ia and (86) it follows that the infinite system (83) possesses in the domain

$$(87)\qquad |\,x\,| < \min\,(1/M, \boldsymbol{a}) = \delta,$$

one and only one solution $y_j(x)$. This solution is holomorphic in the domain (87) and in this domain satisfies the inequalities

$$(88)\qquad |\,y_j(x)\,| < 1, \quad y_j(0) = 0; \qquad (j = 0, \pm 1, \pm 2, \cdots).$$

Accordingly from (53) and (88) the series (54) for $\zeta(\tau)$ is convergent for $|\,\mu\,| < \delta^4$ and $\zeta(\tau)$ converges identically to zero as μ converges to zero.

In order to show that the solution (54) is real for $\mu > 0$ it is only necessary to write (54) in terms of sines and cosines i. e.

$$(54')\qquad \zeta(x) = b_0 + \sum_{\nu=0}^{\infty} (a_\nu \sin \nu\tau + b_\nu \cos \nu\tau),$$

(where the coefficients a_ν and b_ν are real) and to obtain by substitution in the differential equation (51) a real infinite system of conditions for the a_ν and b_ν. That a unique solution of this infinite system exists for which (54′) converges (becoming identically zero for $\mu = 0$) follows since (54′) is only formally different from (54). Furthermore since the infinite system for the a_ν and b_ν is real it follows from the existence theorem on infinite systems stated above that this infinite system possesses a real solution for the a_ν and b_ν. The existence of a real solution (54) is accordingly demonstrated.

5. *The critical case,* $m \equiv 0 \pmod 1$. Let m be a fixed integer.† The infinite system then takes the form

$$(89a)\quad y_j = x\, f_j(x;\, y_0, y_1, y_{-1}, \cdots); \quad (j \neq \pm(m+1), j = 0, \pm 1, \pm 2, \cdots),$$

$$(89b)\quad 0 = A_{100}^{(m+1)}(m) + \Theta_{(m+1)}(x;\, y_0, y_1, y_{-1}, \cdots),$$

† Where we have excluded $m = 0, -1, \pm\infty$ [cf. (48)].

(89c) $\quad 0 = A_{100}^{-(m+1)}(m) + \Theta_{-(m+1)}(x;\, y_0, y_1, y_{-1}, \cdots),$

where we have put

(90) $$\Theta_{\pm(m+1)} \equiv \psi_{\pm(m+1)} - A_{100}^{\pm(m+1)}(m).$$

It follows from (62), (50) and (90) that there exist two positive numbers a and M_3 so that

(91) $$\bar{\Theta}_{\pm(m+1)}(a;\, 1, 1, \cdots) < M_3.$$

The power series $\Theta_{\pm(m+1)}$ in infinitely many variables are therefore uniformly convergent in the domain of infinitely many dimensions

(92) $$|x| < a, \qquad |y_j| < 1; \qquad (j = 0, \pm 1, \pm 2, \cdots).$$

Moreover it follows from the definition of $\Theta_{\pm(m+1)}$ and (54), (55), (57′) and (52) that in the domain of infinitely many dimensions

(93) $$|y_j| < 1; \qquad (j = 0, \pm 1, \pm 2, \cdots),$$

we have

(94) $$\Theta_{\pm(m+1)}(0;\, y_0, y_1, y_{-1}, \cdots) \equiv 0.$$

If we now regard the variables y_{m+1} and $y_{-(m+1)}$ in the infinite system (89a) as parameters λ_1 and λ_2 it follows from (86) by Theorem I (§ 4) that in the domain

(95) $$|x| < \min(1/M, a), \qquad |y_{(m+1)}| < 1, \qquad |y_{-(m+1)}| < 1,$$

the system (89a) possesses one and only one solution

(96) $$y_j = y_j(x;\, y_{(m+1)}, y_{-(m+1)}); \quad (j \neq \pm(m+1),\ j = 0, \pm 1, \pm 2, \cdots).$$

This solution is holomorphic and fulfills in the domain (95) the inequalities

(97) $$|y_j(x;\, y_{(m+1)}, y_{-(m+1)}| < 1 \quad (j \neq \pm(m+1),\ j = 0, \pm 1, \pm 2, \cdots).$$

It follows that the functions

(98) $$\theta_{\pm(m+1)}(x, y_{(m+1)}, y_{-(m+1)}) \equiv$$
$$\Theta_{\pm(m+1)}(x;\, y_0(x;\, y_{(m+1)}, y_{-(m+1)}), y_1(x;\, y_{(m+1)}, y_{-(m+1)}), y_{-1}(x;\, y_{(m+1)}, y_{(m+1)}), \cdots)$$

are holomorphic in the domain (95) of the three complex variables $y_{(m+1)}$ $y_{-(m+1)}$ and x, the power series $\Theta_{\pm(m+1)}$ in the infinitely many variables being uniformly convergent in the domain of infinitely many dimensions (92).

We can now write the remaining conditions (89b) and (89c) in the form

(89b′) $$A_{100}^{(m+1)}(m) + \theta_{(m+1)}(x;\, y_{(m+1)}, y_{-(m+1)}) = 0,$$

(89c′) $$A_{100}^{-(m+1)}(m) + \theta_{-(m+1)}(x;\, y_{(m+1)}, y_{-(m+1)}) = 0.$$

These are the so-called Verzweigungsgleichungen † and (if there exists a solution) serve to determine $y_{(m+1)}$ and $y_{-(m+1)}$ as functions of x. The remaining y_j can then, from (96), be expressed as functions of x; conversely if these equations possess no solution for $y_{(m+1)}$ and $y_{-(m+1)}$ there exists no solution of the differential equation (51) of the desired character. Now it follows from (94) and (98) that $\theta_{\pm(m+1)}(0, y_{(m+1)}, y_{(m+1)}) \equiv 0$. Accordingly if the integer m is such that the constants $A_{100}^{\pm(m+1)}(m)$ in (89b′), (89c′) are not zero there exists for sufficiently small values of x no solution of (89b′), (89c′) and hence no periodic solution of the desired character.

From (50) and (35) it is clear that $A_{100}^{\pm(m+1)}(m)$ are linear combinations, the coefficients of which are functions of m, of the coefficients of Laplace (which themselves are functions of m). Accordingly, for a given value of m, the numerical values of $A_{100}^{\pm(m+1)}(m)$ can be calculated from tables of the Laplace coefficients.‡ The author has calculated § these coefficients for $|m| \leqq 4$ and obtains

(99) $$A_{100}^{\pm(m+1)}(m) \neq 0 \qquad (m = 1, \pm 2, \pm 3, \pm 4),$$

which is sufficient to show that for sufficiently small values of x the equations (89b′) (89c′) possess no solution for $y_{(m+1)}$ and $y_{-(m+1)}$.

THE JOHNS HOPKINS UNIVERSITY.

---

† Cf. footnote †, p. 260.

‡ Tables giving the numerical values for the coefficients of Laplace for various values of a (and therefore of m) are given for instance by Runkle. Not having access to these tables, recourse was had to a memoir of Leverrier in the second volume of the *Annales de l'Observatoire de Paris*, in which he gives recursion formulas for the coefficients of Laplace and their well known expansion in a hypergeometric series according to powers of a, in a form suitable for calculation.

§ The inequality (99) can be established for very large values of m by means of well known asymptotic methods.

3

NOTE ON THE CONSTANTS OF THE DISTURBING FUNCTION.

By K. P. Williams.

The method for obtaining the quantities $c_n^{(i,j)}$ that appear in Newcomb's development of the disturbing function that was given in this journal recently * had in view primarily the case where the ratio $s = a/a'$ is large. For small values of s the work can be greatly shortened by using the following method for finding the $c_1^{(i,j)}$, $j > 0$.

Making use of (16) and the series (19) for H, we have

$$\frac{c_1^{(i,j+1)}}{c_1^{(i,j)}} = \frac{2(2j+1)(2j+2i+1)}{j+i+1}\left(\frac{s}{2}\right)^2 [1 + b_1 y + \cdots].$$

When the value for a_1 is used we find

$$b_1 = -1 + \frac{1+2i}{4(i+j+1)(i+j+2)}.$$

Hence

$$\frac{c_1^{(i,j+1)}}{c_1^{(i,j)}} = s^2(2j+1)M,$$

where M is a number not much different from unity.

If s is small we can easily observe by this formula, after the $c_1^{(i)}$ have been found, the lowest element $c_1^{(k,1)}$ in the second column that is not negligible. Putting $c_1^{(k+1,1)} = 0$ we have from (45)

$$c_1^{(k,1)} = sc_1^{(k+1)}.$$

The column can be completed as before. Similar remarks apply to succeding columns.

The following errata appear in the numerical series given at the end of the article cited:

The numerical coefficient in $c_1^{(9)}$ should be $2 \cdot 5 \cdot 11 \cdot 13 \cdot 17$; that of $c_1^{(10,3)}$ should be $1280 \cdot 17 \cdot 19 \cdot 21 \cdot 23 \cdot 25$.

Indiana University,
Bloomington, Indiana.

* K. P. Williams, "The Constants of the Disturbing Function," *American Journal of Mathematics*, Vol. 52, pp. 571-584. References are to this article.

CONDUCTORS IN AN ELECTROMAGNETIC FIELD (E^0e^{pt}, H^0e^{pt}).

By F. H. Murray.

The physical problem of a system of conductors in an impressed electromagnetic field of the form ($\mathbf{E}^0e^{pt}$, $\mathbf{H}^0e^{pt}$) leads to the well-known field equations in space and in the conductors, with boundary conditions at their surfaces. It is proposed here to develop certain formulas required in the mathematical discussion of the general problem; by means of certain vector identities it is shown that the field which must be added to the impressed field exterior to the conductors can be represented by integrals over the surfaces of the conductors, which are equivalent to a representation in terms of surface distributions of electric and magnetic doublets.* Part I is devoted to the derivation and discussion of these identities; in Part II a discussion is given of the special case of perfect conductors. The representation leads to a system of integral equations which can be reduced to the type of Fredholm, and it is shown that the exceptional values of p, for which the equations corresponding to an arbitrary impressed field do not possess a solution, must be the values which correspond to cavity-radiation for at least one conductor. While the discussion of certain points could be abbreviated by appeals to physical intuition, it appeared desirable, for possible applications to high frequencies, to give formal proofs of all propositions stated.

From the general formulas of Part I can be derived the equations for the case of wire conductors, in a high-frequency radiation field; a discussion of these equations and applications will be given in another paper.

Part I. General Formulas.

1. Let the system of surfaces $S_1, S_2 \cdots S_n$ be denoted by S; each surface is assumed to have a continuous tangent plane and to be such that if an arbitrary point (x_0, y_0, z_0) on S is given, a transformation of the coördinate axes can be made such that all points of S in some neighborhood of this point can be represented in the form $z = f(x, y)$, where f possesses continuous partial derivatives of the first three orders, and

$$0 = f(0,0) = \partial f/\partial x \,\Big|_{x=y=0} = \partial f/\partial y \,\Big|_{x=y=0}.$$

* Related formulae have been developed by Hasenörl, *Physikalische Zeitschrift*, Band 7 (1906), p. 37. For the special case of perfect conductors see MacDonald, *Electric Waves*, p. 15, also *Proceedings of the London Mathematical Society* (2), Vol. 10 (1911), p. 91.

Let μ, σ, ϵ denote the permeability, conductivity, and specific inductive capacity, respectively, of the medium exterior to S (exterior to every S_i), while μ', σ', ϵ' denote the constants of the interior, fixed for any conductor, but not necessarily the same for different conductors.

If the impressed and added fields are written in the form

$$\begin{array}{ll} \boldsymbol{E}^0(xyz)e^{pt}, & \boldsymbol{H}^0(xyz)e^{pt} \\ \boldsymbol{E}\ (xyz)e^{pt}, & \boldsymbol{H}\ (xyz)e^{pt} \end{array}$$

and if

$$\lambda = 1/c(4\pi\sigma + \epsilon p), \quad \kappa = -\mu p/c, \quad h = (-\kappa\lambda)^{1/2}$$

the equations of Maxwell become,

$$(1.1) \qquad \begin{array}{ll} \lambda \boldsymbol{E} = \operatorname{curl} \boldsymbol{H} & \lambda' \boldsymbol{E}' = \operatorname{curl} \boldsymbol{H}' \\ \kappa \boldsymbol{H} = \operatorname{curl} \boldsymbol{E} & \kappa' \boldsymbol{H}' = \operatorname{curl} \boldsymbol{E}' \end{array}$$

primes denoting interior values. The components of $\boldsymbol{E}$, $\boldsymbol{H}$ satisfy the wave equation

$$(1.2) \qquad (\Delta - h^2)\phi = (\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2 - h^2)\phi = 0.$$

This equation has the fundamental solution

$$\phi = (1/r)e^{-hr}, \qquad r = [(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2]^{1/2}.$$

The real part of h is assumed positive or zero; it is at once seen that if p lies on the right of the imaginary axis in its complex plane, a branch of h which is positive when p is positive real will have its real part positive and this branch of the function will be used throughout. The boundary conditions on S are that the tangential components of $\boldsymbol{E}^0 + \boldsymbol{E}$ are equal to the corresponding tangential components of $\boldsymbol{E}'$, similarly, $(\boldsymbol{H}^0 + \boldsymbol{H})_{\text{tang.}} = \boldsymbol{H}'_{\text{tang.}}$.

2. Let Σ be the surface of a large sphere enclosing S, while γ is the surface of a small sphere enclosing a point $(x_0y_0z_0)$ exterior to S, but interior to Σ; if $\boldsymbol{u}$, $\boldsymbol{v}$ are two vectors which are continuous, with their partial derivatives of the first two orders in the interior of Σ, exterior to S and γ, we have the identity *

$$[\boldsymbol{u}, \operatorname{curl} \boldsymbol{v}]_n + \boldsymbol{u}_n \operatorname{div} \boldsymbol{v} - [\boldsymbol{v}, \operatorname{curl} \boldsymbol{u}]_n \\ - \boldsymbol{v}_n \operatorname{div} \boldsymbol{u} = \boldsymbol{u} d\boldsymbol{v}/dn - \boldsymbol{v} d\boldsymbol{u}/dn + \operatorname{curl}_n [\boldsymbol{u}, \boldsymbol{v}].$$

From Green's formula in vector notation

$$\iiint_{(S+\Sigma+\gamma)} \{\boldsymbol{u}\Delta\boldsymbol{v} - \boldsymbol{v}\Delta\boldsymbol{u}\}\, dx\, dy\, dz = -\iint_{S+\Sigma+\gamma} (\boldsymbol{u} d\boldsymbol{v}/dn - \boldsymbol{v} d\boldsymbol{u}/dn)\, dS.$$

* Weyl, "Über die Randwertaufgabe der Strahlungstheorie und asymptotische Spectralgesetze, *Crelle's Journal*, Band 143 (1913), p. 182.

Also,

$$0 = \iint_{S+\Sigma+\gamma} \operatorname{curl}_n [\boldsymbol{u}, \boldsymbol{v}]\, dS$$

consequently

$$\iiint_{(S+\Sigma+\gamma)} \{\boldsymbol{u}\Delta\boldsymbol{v} - \boldsymbol{v}\Delta\boldsymbol{u}\}\, dx\, dy\, dz = \iint_{S+\Sigma+\gamma} \left\{ \begin{array}{l} -[\boldsymbol{u}, \operatorname{curl} \boldsymbol{v}]_n - \boldsymbol{u}_n \operatorname{div} \boldsymbol{v} \\ +[\boldsymbol{v}, \operatorname{curl} \boldsymbol{u}]_n + \boldsymbol{v}_n \operatorname{div} \boldsymbol{u} \end{array} \right\} dS.$$

If the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are solutions of the equation (1. 2), the integral on the left vanishes; if in addition the components of $\boldsymbol{u}$ have continuous partial derivatives of the first two orders interior to γ, while $\boldsymbol{v}$ is one of the vectors

$$\boldsymbol{v}' = (\phi, 0, 0), \quad \boldsymbol{v}'' = (0, \phi, 0), \quad \boldsymbol{v}''' = (0, 0, \phi),$$

from Stokes' theorem

$$\iint_{\gamma} \operatorname{curl}_n [\boldsymbol{u}, \boldsymbol{v}]\, dS = 0.$$

Consequently, if the vectors above are substituted in succession in the identity

$$-\iint_{\gamma} \operatorname{curl}_n [\boldsymbol{u}, \boldsymbol{v}]\, dS - \iint_{\gamma} \left(\boldsymbol{u}\frac{d\boldsymbol{v}}{dn} - \boldsymbol{v}\frac{d\boldsymbol{u}}{dn}\right) dS$$
$$= \iint_{S+\Sigma} \left\{ \begin{array}{l} [\boldsymbol{u}, \operatorname{curl} \boldsymbol{v}]_n + \boldsymbol{u}_n \operatorname{div} \mathbf{v} \\ -[\boldsymbol{v}, \operatorname{curl} \boldsymbol{u}]_n - \boldsymbol{v}_n \operatorname{div} \mathbf{u} \end{array} \right\} dS$$

and the radius of γ is made to approach zero, the result is,

$$4w\boldsymbol{u}_x = \iint_{S+\Sigma} \{[\boldsymbol{n}, \boldsymbol{u}] \cdot \operatorname{curl} \boldsymbol{v}' + \boldsymbol{u}_n \operatorname{div} \boldsymbol{v}' - [\boldsymbol{n}, \boldsymbol{v}'] \operatorname{curl} \boldsymbol{u} - \boldsymbol{v}_n' \operatorname{div} \boldsymbol{u}\} dS.$$

The components $\boldsymbol{u}_y$, $\boldsymbol{u}_z$ have the same representation in terms of $\boldsymbol{v}''$, $\boldsymbol{v}'''$, respectively.

If $\boldsymbol{u} = \boldsymbol{E}$, since $\operatorname{div} \boldsymbol{E} = 0$, one obtains from the field equations

$$4\pi\boldsymbol{E}_x = \iint_{S+\Sigma} \{[\boldsymbol{n}, \boldsymbol{E}] \cdot \operatorname{curl} \boldsymbol{v}' + \boldsymbol{E}_n \operatorname{div} \boldsymbol{v}' + \kappa[\boldsymbol{n}, \boldsymbol{H}]\boldsymbol{v}'\} dS.$$

If the real part of h is not zero, the integral over Σ approaches zero if $\boldsymbol{E}$ is bounded at infinity; if the real part of h is zero, but $h \neq 0$, while $(\boldsymbol{E}, \boldsymbol{H})$ behaves at infinity like the field of a system of diverging spherical waves

$$\lim_{R\to\infty} Re^{hR}\boldsymbol{E} = \boldsymbol{E}_1(\theta, \phi), \quad \boldsymbol{E} \sim \boldsymbol{E}_1(\theta, \phi)\,(1/R)\,e^{-hR}$$

and if the approach to the limiting value is uniform with respect to the polar angles θ, ϕ,

$$\iint_{\Sigma} \{[\boldsymbol{n}, \boldsymbol{E}] \operatorname{curl} \boldsymbol{v}' + \boldsymbol{E}_n \operatorname{div} \boldsymbol{v}' + \kappa[\boldsymbol{n}, \boldsymbol{H}] \cdot \boldsymbol{v}'\} dS \sim Ce^{-2hR}.$$

Consequently, if $C \neq 0$, $4\pi \boldsymbol{E}(x_0, y_0, z_0)$ does not approach a limit as $R \to \infty$. But since $\boldsymbol{E}$ is independent of R, $C = 0$.
Hence

$$(1.3)\quad \begin{aligned} 4\pi \boldsymbol{E}_x &= \iint_S \{[\boldsymbol{n}, \boldsymbol{E}] \cdot \operatorname{curl} \boldsymbol{v}' + \boldsymbol{E}_n \operatorname{div} \boldsymbol{v}' + \kappa[\boldsymbol{n}, \boldsymbol{H}]\boldsymbol{v}'\} dS \\ 4\pi \boldsymbol{E}_y &= \iint_S \{[\boldsymbol{n}, \boldsymbol{E}] \cdot \operatorname{curl} \boldsymbol{v}'' + \boldsymbol{E}_n \operatorname{div} \boldsymbol{v}'' + \kappa[\boldsymbol{n}, \boldsymbol{H}]\boldsymbol{v}''\} dS \\ 4\pi \boldsymbol{E}_z &= \iint_S \{[\boldsymbol{n}, \boldsymbol{E}] \cdot \operatorname{curl} \boldsymbol{v}''' + \boldsymbol{E}_n \operatorname{div} \boldsymbol{v}''' + \kappa[\boldsymbol{n}, \boldsymbol{H}]\mathrm{v}'''\} dS. \end{aligned}$$

The corresponding representation of $\boldsymbol{H}$ is obtained by interchanging $\boldsymbol{E}$ and $\boldsymbol{H}$ everywhere, while κ is replaced by λ.

If the definitions of $\boldsymbol{v}', \boldsymbol{v}'', \boldsymbol{v}'''$ are employed, the result is,

$$(1.4)\quad \begin{aligned} 4\pi \boldsymbol{E} &= \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{E}]\, \phi dS - \operatorname{grad} \iint_S \boldsymbol{E}_n \phi dS + \kappa \iint_S [\boldsymbol{n}, \boldsymbol{H}]\, \phi dS \\ 4\pi \boldsymbol{H} &= \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{H}]\, \phi dS - \operatorname{grad} \iint_S \boldsymbol{H}_n \phi dS + \lambda \iint_S [\boldsymbol{n}, \boldsymbol{E}]\, \phi dS. \end{aligned}$$

If the point (x_0, y_0, z_0) had been taken interior to any S_i, the integral over γ could have been omitted in the discussion, and the left-hand side of (1.4) would be zero. The normal indicated above is the exterior normal; the formulas for the interior of any surface S_i are the same as above, with $\boldsymbol{n}$ indicating the interior normal, S replaced by S_i, and ϕ replaced by ϕ' constructed with the constants of the conductor.

The impressed field $(\boldsymbol{E}^0, \boldsymbol{H}^0)$ has no singularities interior to any S_i, hence if (x, y, z) is exterior to S_i,

$$0 = \operatorname{curl} \iint_{S_i} [\boldsymbol{n}', \boldsymbol{E}^0]\, \phi dS - \operatorname{grad} \iint_{S_i} \boldsymbol{E}_{n'}{}^0 \phi dS + \kappa \iint_{S_i} [\boldsymbol{n}', \boldsymbol{H}^0]\, \phi dS.$$

Replacing $\boldsymbol{n}'$ by $-\boldsymbol{n}$, and observing that the tangential components of $\boldsymbol{E}^0$, $\boldsymbol{H}^0$ are the same on both sides of S_i, $\boldsymbol{E}_{n'}{}^0 = -\boldsymbol{E}_n{}^0$, this can be written

$$0 = \operatorname{curl} \iint_{S_i} [\boldsymbol{n}, \boldsymbol{E}^0]\, \phi dS - \operatorname{grad} \iint_{S_i} \boldsymbol{E}_n{}^0 \phi dS + \kappa \iint_{S_i} [\boldsymbol{n}, \boldsymbol{H}^0]\, \phi dS.$$

Summing with respect to i and adding the result to the first equation (1.4), and repeating the process for the second equation,

(1.5)

$$\begin{aligned} 4\pi \boldsymbol{E} &= \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{E} + \boldsymbol{E}^0] \phi dS - \operatorname{grad} \iint_S (\boldsymbol{E}_n + \boldsymbol{E}_n{}^0) \phi dS + \kappa \iint_S [\boldsymbol{n}, \boldsymbol{H} + \boldsymbol{H}^0] \phi dS \\ 4\pi \boldsymbol{H} &= \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{H} + \boldsymbol{H}^0] \phi dS - \operatorname{grad} \iint_S (\boldsymbol{H}_n + \boldsymbol{H}_n{}^0) \phi dS + \lambda \iint_S [\boldsymbol{n}, \boldsymbol{E} + \boldsymbol{E}^0] \phi dS. \end{aligned}$$

From the boundary conditions

$$[\boldsymbol{n}, \boldsymbol{E} + \boldsymbol{E}^0] = [\boldsymbol{n}, \boldsymbol{E}'], \quad [\boldsymbol{n}, \boldsymbol{H} + \boldsymbol{H}^0] = [\boldsymbol{n}, \boldsymbol{H}'],$$

and the derived conditions *

$$\lambda(\boldsymbol{E}_n{}^0 + \boldsymbol{E}_n) = \lambda' \boldsymbol{E}_n', \quad \mu(\boldsymbol{H}_n{}^0 + \boldsymbol{H}_n) = \mu' \boldsymbol{H}_n'$$

and (1. 5) becomes

$$(1.6)\quad \begin{aligned} 4\pi \boldsymbol{E} &= \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{E}'] \phi dS - \operatorname{grad} \iint_S (\lambda'/\lambda) \boldsymbol{E}_n' \phi dS + \kappa \iint_S [\boldsymbol{n}, \boldsymbol{H}'] \phi dS \\ 4\pi \boldsymbol{H} &= \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{H}'] \phi dS - \operatorname{grad} \iint_S (\kappa'/\kappa) \boldsymbol{H}_n' \phi dS + \lambda \iint_S [\boldsymbol{n}, \boldsymbol{E}'] \phi dS. \end{aligned}$$

3. If (x_1, y_1, z_1) represents a point on S in the formulas (1. 4), while (x, y, z) denotes a point exterior to S, and if $\boldsymbol{J} = [\boldsymbol{n}, \boldsymbol{H}]$, the condition $\operatorname{div} \boldsymbol{E} = 0$ becomes,

$$-\iint_S \boldsymbol{E}_n \Delta \phi dS + \kappa \iint_S (\boldsymbol{J}_{x_1} \partial\phi/\partial x + \boldsymbol{J}_{y_1} \partial\phi/\partial y + \boldsymbol{J}_{z_1} \partial\phi/\partial z) dS = 0$$

or since $h^2 = -\kappa\lambda$,

$$\iint_S (-\lambda \boldsymbol{E}_n \phi + \boldsymbol{J}_{x_1} \partial\phi/\partial x_1 + \boldsymbol{J}_{y_1} \partial\phi/\partial y_1 + \boldsymbol{J}_{z_1} \partial\phi/\partial z_1) dS = 0.$$

On the surface S let (u, v) be isothermal Gaussian coördinates for a small part of the surface Ω bounded by an ordinary curve C; if $\phi_x = \partial\phi/\partial x$ etc., while (u, v, n) forms a right-handed system; then

$$\begin{matrix} \alpha' = \partial x_1/\partial u & \alpha'' = \partial x_1/\partial v & l = \partial x_1/\partial n \,|_{n=0} \\ \beta' = \partial y_1/\partial u & \beta'' = \partial y_1/\partial v & m = \partial y_1/\partial n \,|_{n=0} \\ \gamma' = \partial z_1/\partial u & \gamma'' = \partial z_1/\partial v & n = \partial z_1/\partial n \,|_{n=0} \end{matrix} \qquad \begin{vmatrix} \alpha' & \alpha'' & l \\ \beta' & \beta'' & m \\ \gamma' & \gamma'' & n \end{vmatrix} = 1$$

$$\begin{aligned} \boldsymbol{H}_u &= \alpha' \boldsymbol{H}_{x_1} + \beta' \boldsymbol{H}_{y_1} + \gamma' \boldsymbol{H}_{z_1}, & \boldsymbol{H}_{x_1} &= \alpha' \boldsymbol{H}_u + \beta'' \boldsymbol{H}_v + l \boldsymbol{H}_n \\ \boldsymbol{H}_v &= \alpha'' \boldsymbol{H}_{x_1} + \beta'' \boldsymbol{H}_{y_1} + \gamma'' \boldsymbol{H}_{z_1}, & \boldsymbol{H}_{y_1} &= \beta' \boldsymbol{H}_u + \alpha'' \boldsymbol{H}_v + m \boldsymbol{H}_n \\ \boldsymbol{H}_n &= l \boldsymbol{H}_{x_1} + m \boldsymbol{H}_{y_1} + n \boldsymbol{H}_{z_1}, & \boldsymbol{H}_{z_1} &= \gamma' \boldsymbol{H}_u + \gamma'' \boldsymbol{H}_v + n \boldsymbol{H}_n. \end{aligned}$$

Then

$$\begin{aligned} \boldsymbol{J}_{x_1} \partial\phi/\partial x_1 + \boldsymbol{J}_{y_1} \partial\phi/\partial y_1 + \boldsymbol{J}_{z_1} \partial\phi/\partial z_1 &= \boldsymbol{H}_u \partial\phi/\partial v - \boldsymbol{H}_v \partial\phi/\partial u. \\ &= -\partial/\partial u (\boldsymbol{H}_v \phi) + \partial/\partial v (\boldsymbol{H}_u \phi) + \phi\, [\partial \boldsymbol{H}_v/\partial u - \partial \boldsymbol{H}_u/\partial v]. \end{aligned}$$

Consequently

$$\begin{aligned} &\iint_\Omega (\boldsymbol{J}_{x_1} \partial\phi/\partial x_1 + \boldsymbol{J}_{y_1} \partial\phi/\partial y_1 + \boldsymbol{J}_{z_1} \partial\phi/\partial z_1) dS \\ &= \iint_\Omega [\partial/\partial v (\boldsymbol{H}_u \phi) - \partial/\partial u (\boldsymbol{H}_v \phi)]\, du\, dv + \iint_\Omega \phi (\partial \boldsymbol{H}_v/\partial u - \partial \boldsymbol{H}_u/\partial v)\, du\, dv \\ &= -\int_C \phi\, [\boldsymbol{H}_u du + \boldsymbol{H}_v dv] + \iint_\Omega \phi (\operatorname{curl} \boldsymbol{H})_n dS \end{aligned}$$

---

* These equations result from the fact that the identity $\boldsymbol{E}_t' = (\boldsymbol{E} + \boldsymbol{E}^0)_t$ implies that an identity is obtained by differentiating each side tangentially; the normal component of $(\boldsymbol{H} + \boldsymbol{H}^0)$ is expressed in terms of the tangential derivatives of $\boldsymbol{E} + \boldsymbol{E}^0$, hence the second relation. The first is obtained in the same manner.

from Stokes' theorem in the (u, v) plane. Now if each surface S_i is divided up into a number of parts Ω, each bounding arc of the curves C occurs twice, the integration being in opposite directions, while the integrand of the line integrals is independent of the choice of Gaussian coördinates; the sum of the integrals over the curves C vanishes. Hence

$$\iint_{S_i} (\boldsymbol{J}_{x_1}\partial\phi/\partial x_1 + \boldsymbol{J}_{y_1}\partial\phi/\partial y_1 + \boldsymbol{J}_{z_1}\partial\phi/\partial z_1)\,dS = \iint_{S_i} \phi(\operatorname{curl}\boldsymbol{H})_n dS = \iint_{S_i} \lambda \boldsymbol{E}_n \phi dS.$$

Hence $\operatorname{div}\boldsymbol{E} = 0$; similarly, $\operatorname{div}\boldsymbol{H} = 0$.

A representation of the field $(\boldsymbol{E}, \boldsymbol{H})$ in terms of surface distributions of electric and magnetic doublets results immediately. Let

$$4\pi\boldsymbol{H}_1 = \operatorname{curl}\iint_S [\boldsymbol{n}, \boldsymbol{H}]\,\phi dS, \qquad \Phi = \iint_S \boldsymbol{E}_n \phi dS.$$

The electric field corresponding to $\boldsymbol{H}_1$ results from the field equations:

$$\boldsymbol{E}_1 = 1/\lambda \operatorname{curl}\boldsymbol{H}_1 = (1/4\pi\lambda)\{\operatorname{grad}\operatorname{div}\iint_S [\boldsymbol{n}, \boldsymbol{H}]\,\phi dS - h^2 \iint_S [\boldsymbol{n}, \boldsymbol{H}]\,\phi dS\}.$$

Since

$$\operatorname{div}\{-\operatorname{grad}\Phi + \kappa\iint_S [\boldsymbol{n}, \boldsymbol{H}]\,\phi dS\} = 0,$$

we have

$$h^2\Phi = \operatorname{div}\kappa\iint_S [\boldsymbol{n}, \boldsymbol{H}]\,\phi dS. \tag{3.1}$$

Hence

$$4\pi\boldsymbol{E}_1 = -\operatorname{grad}\Phi + \kappa\iint_S [\boldsymbol{n}, \boldsymbol{H}]\,\phi dS.$$

Similarly, if

$$\psi = \iint_S \boldsymbol{H}_n \phi dS,$$

the vectors

$$4\pi\boldsymbol{E}_2 = \operatorname{curl}\iint_S [\boldsymbol{n}, \boldsymbol{E}]\,\phi dS,$$
$$4\pi\boldsymbol{H}_2 = -\operatorname{grad}\psi + \lambda\iint_S [\boldsymbol{n}, \boldsymbol{H}]\,\phi dS$$

form a system satisfying the field equations; since

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2, \quad \boldsymbol{H} = \boldsymbol{H}_1 + \boldsymbol{H}_2$$

it follows that an arbitrary diverging field $(\boldsymbol{E}, \boldsymbol{H})$ can be represented as the sum of surface distributions of electric and magnetic doublets.

4. Another representation of the field $(\boldsymbol{E}, \boldsymbol{H})$ in terms of surface and volume integrals, which reduces to a well-known representation when h^2 is neglected in comparison with h'^2, can be obtained as follows. As before,

let $(\boldsymbol{E}',\boldsymbol{H}')$ denote the field interior to any conductor; if $\boldsymbol{v}$ is one of the vectors $\boldsymbol{v}', \boldsymbol{v}'', \boldsymbol{v}'''$, an identity already used gives the equation, in which $\boldsymbol{n}'$ is the interior normal,

$$\iiint_{(S_i)\text{int.}} (\boldsymbol{E}'\Delta\boldsymbol{v} - \boldsymbol{v}\Delta\boldsymbol{E}')\,dx_1dy_1dz_1$$
$$= - \iint_{S_i} \{[\boldsymbol{n}',\boldsymbol{E}'] \operatorname{curl} \boldsymbol{v} + \boldsymbol{E}'_{n'} \operatorname{div} \boldsymbol{v} - [\boldsymbol{n}',\boldsymbol{v}] \operatorname{curl} \boldsymbol{E}'\}dS.$$

Since
$$\Delta\boldsymbol{v} = h^2\boldsymbol{v}, \quad \Delta\boldsymbol{E}' = h'^2\boldsymbol{E}',$$
this becomes

$$\iiint_{(S_i)\text{int.}} (h^2 - h'^2)\boldsymbol{E}'\boldsymbol{v}\,dx_1dy_1dz_1$$
$$= - \iint_{S_i} \{[\boldsymbol{n}',\boldsymbol{E}'] \operatorname{curl} \boldsymbol{v} + \boldsymbol{E}'_{n'} \operatorname{div} \boldsymbol{v} + \kappa' [\boldsymbol{n}',\boldsymbol{H}']\, \boldsymbol{v}\}dS.$$

Replacing $\boldsymbol{n}'$ by $-\boldsymbol{n}$, and summing with respect to all conductors,

$$\iiint_{(S)\text{int.}} (h^2 - h'^2)\boldsymbol{E}'\boldsymbol{v}\,dx_1dy_1dz_1$$
$$= \iint_{S} \{[\boldsymbol{n},\boldsymbol{E}'] \operatorname{curl} \boldsymbol{v} + \boldsymbol{E}_n' \operatorname{div} \boldsymbol{v} + \kappa' [\boldsymbol{n},\boldsymbol{H}']\, \mathrm{v}\}dS$$

which is easily transformed into

$$\iiint_{(S)\text{int.}} (h^2 - h'^2)\boldsymbol{E}'\phi\, dx_1dy_1dz_1$$
$$= \operatorname{curl} \iint_S [\boldsymbol{n},\boldsymbol{E}']\,\phi dS - \operatorname{grad} \iint_S \boldsymbol{E}_n'\phi dS + \iint_S \kappa' [\boldsymbol{n},\boldsymbol{H}']\,\phi dS.$$

Combining this identity with the first equation of (1.6), one obtains

$$(4.1) \quad 4\pi\boldsymbol{E} = \iiint_{(S)\text{int.}} (h^2 - h'^2)\boldsymbol{E}'\phi dx_1dy_1dz_1$$
$$- \operatorname{grad} \iint_S [(\lambda' - \lambda)/\lambda]\, \boldsymbol{E}_n'\phi dS + \iint_S (\kappa - \kappa')[\boldsymbol{n},\boldsymbol{H}']\,\phi dS.$$

From the field equations

$$(4.2) \quad 4\pi\boldsymbol{H} = \operatorname{curl} \iiint_{(S)\text{int.}} [(h^2 - h'^2)/\kappa]\, \boldsymbol{E}'\phi dx_1dy_1dz_1$$
$$+ \operatorname{curl} \iint_S [(\kappa - \kappa')/\kappa]\, [\boldsymbol{n},\boldsymbol{H}']\,\phi dS.$$

These formulas bring into evidence the vanishing of the field $(\boldsymbol{E},\boldsymbol{H})$ if $\lambda = \lambda', \kappa = \kappa'$.

PART II. PERFECT CONDUCTORS.

1. If all the surfaces S_i are assumed to bound perfect conductors, on each $(\boldsymbol{E}^0 + \boldsymbol{E})_{\text{tang.}} = 0$; the representation of the magnetic field * reduces to

* If $\sigma \neq 0$ exterior to S, another treatment of the problem is possible; see "The Electromagnetic Field Exterior to a System of Perfectly Reflecting Surfaces," *Proceedings of the National Academy of Sciences*, Vol. 16, No. 5 (May, 1930), pp. 353-357.

$$4\pi \boldsymbol{H} = \text{curl} \iint_S [\boldsymbol{n}, \boldsymbol{H}'] \, \phi dS.$$

Let $\boldsymbol{J} = [\boldsymbol{n}, \boldsymbol{H}']$. Introducing the isothermal coördinate system (u, v) on the surface, and the direction cosines of the u, v curves $(\alpha', \beta', \gamma')$, $(\alpha'', \beta'', \gamma'')$, respectively, let components at a fixed point be denoted by the subscript 0; then if (x, y, z) is a point on the normal at (x_0, y_0, z_0), we have

$$4\pi(\alpha_0'\boldsymbol{H}_x + \beta_0'\boldsymbol{H}_y + \gamma_0'\boldsymbol{H}_z) = -\iint_S \begin{vmatrix} \alpha_0' & \beta_0' & \gamma_0' \\ \partial\phi/\partial x_1 & \partial\phi/\partial y_1 & \partial\phi/\partial z_1 \\ \boldsymbol{J}_{x_1} & \boldsymbol{J}_{y_1} & \boldsymbol{J}_{z_1} \end{vmatrix} dS.$$

On the surface S,

$$\begin{aligned} \partial\phi/\partial x &= \alpha'\partial\phi/\partial u + \alpha''\partial\phi/\partial v + l\partial\phi/\partial n \\ \partial\phi/\partial y &= \beta'\partial\phi/\partial u + \beta''\partial\phi/\partial v + m\partial\phi/\partial n \\ \partial\phi/\partial z &= \gamma'\partial\phi/\partial u + \gamma''\partial\phi/\partial v + n\partial\phi/\partial n \end{aligned}$$

consequently

$$\begin{aligned} &4\pi\boldsymbol{H}_{u_0}(u_0, v_0, n) \\ &\quad = -\iint_S \begin{vmatrix} \alpha' & \beta' & \gamma' \\ \alpha''\partial\phi/\partial v + l\partial\phi/\partial n & \beta''\partial\phi/\partial v + m\partial\phi/\partial n & \gamma''\partial\phi/\partial u + n\partial\phi/\partial n \\ \boldsymbol{J}_x & \boldsymbol{J}_y & \boldsymbol{J}_z \end{vmatrix} dS \\ &\quad + \iint_S \{(\alpha_0' - \alpha')(\boldsymbol{J}_z\partial\phi/\partial y - \boldsymbol{J}_y\partial\phi/\partial z) + (\beta_0' - \beta')(\boldsymbol{J}_x\partial\phi/\partial z - \boldsymbol{J}_z\partial\phi/\partial x) \\ &\qquad + (\gamma_0' - \gamma')(\boldsymbol{J}_y\partial\phi/\partial x - \boldsymbol{J}_x\partial\phi/\partial y)\} dS. \end{aligned}$$

Now

$$\boldsymbol{J}_x = \alpha'\boldsymbol{J}_u + \alpha''\boldsymbol{J}_v, \quad \boldsymbol{J}_y = \beta'\boldsymbol{J}_u + \beta''\boldsymbol{J}_v, \quad \boldsymbol{J}_z = \gamma'\boldsymbol{J}_u + \gamma''\boldsymbol{J}_v,$$

from which

$$4\pi\boldsymbol{H}_{u_0}(u_0, v_0, n) = \iint_S (\boldsymbol{J}_v d\phi/dn) dS + \iint_S (A\boldsymbol{J}_u + B\boldsymbol{J}_v) dS.$$

Since

$$\boldsymbol{J}_v = \boldsymbol{H}_u', \qquad \boldsymbol{J}_u = -\boldsymbol{H}_v'$$

the preceding equation becomes,

$$(1.1) \qquad 4\pi\boldsymbol{H}_{u_0}(u_0, v_0, n) = \iint_S (\boldsymbol{H}_u' d\phi/dn) dS + \iint_S (B\boldsymbol{H}_u' - A\boldsymbol{H}_v') dS.$$

Now the function ϕ becomes infinite like $1/r$, and from the theory of the Newtonian potential,

$$\begin{aligned} 4\pi\boldsymbol{H}_{u_0}(u_0, v_0, 0) &= 2\pi\boldsymbol{H}_{u_0}'(u_0, v_0, 0) \\ &\quad + \iint_S \boldsymbol{H}_u'[d\phi/dn] dS + \iint_S (B\boldsymbol{H}_u' - A\boldsymbol{H}_v') dS \end{aligned}$$

since A, B become infinite only to the first order on the surface. This equation may also be written

$$(1.2)\quad 2\pi\boldsymbol{H}_{u_0}{}' = 4\pi\mathbf{H}_{u_0}{}^0 - \iint\limits_S \begin{vmatrix} \alpha_0' & \beta_0' & \gamma_0' \\ [\partial\phi/\partial x]_0 & [\partial\phi/\partial y]_0 & [\partial\phi/\partial z]_0 \\ [\boldsymbol{n}, \boldsymbol{H}']_x & [\boldsymbol{n}, \boldsymbol{H}']_y & [\boldsymbol{n}, \boldsymbol{H}']_z \end{vmatrix} dS$$

the bracket indicating the values on the surface. Similarly,

$$2\pi\boldsymbol{H}_{v_0}{}' = 4\pi\boldsymbol{H}_{v_0}{}^0 - \iint\limits_S \begin{vmatrix} \alpha_0'' & \beta_0'' & \gamma_0'' \\ [\partial\phi/\partial x]_0 & [\partial\phi/\partial y]_0 & [\partial\phi/\partial z]_0 \\ [\boldsymbol{n}, \boldsymbol{H}']_x & [\boldsymbol{n}, \boldsymbol{H}']_y & [\boldsymbol{n}, \boldsymbol{H}']_z \end{vmatrix} dS.$$

These equations form a system of integral equations which can be reduced to the type of Fredholm by iteration, and consequently possess a unique solution unless the homogeneous equations, corresponding to no impressed field, have a solution. In this case the solution of the homogeneous equations defines a field $(\boldsymbol{E}, \boldsymbol{H})$ by means of

$$4\pi\boldsymbol{H} = \text{curl} \iint\limits_S [\boldsymbol{n}, \bar{\boldsymbol{H}}]\, \phi dS,$$

$$\boldsymbol{E} = 1/\lambda \text{ curl } \boldsymbol{H}.$$

The integral equations express that the condition $[\boldsymbol{n}, \boldsymbol{H}] \to [\boldsymbol{n}_0, \bar{\boldsymbol{H}}]$ is satisfied, hence $\boldsymbol{H}$ is expressed in terms of its tangential components. Hence in the equations (1.4), Part I, the first integral representing $\boldsymbol{H}$ is the same. This is only possible if

$$0 = -\text{grad} \iint\limits_S \boldsymbol{H}_n \phi dS + \lambda \iint\limits_S [\boldsymbol{n}, \boldsymbol{E}]\, \phi dS$$

and

$$4\pi\boldsymbol{E} = -\text{grad} \iint\limits_S \boldsymbol{E}_n \phi dS + \kappa \iint\limits_S [\boldsymbol{n}, \boldsymbol{H}]\, \phi dS.$$

The surface integrals which represent $(\boldsymbol{E}, \boldsymbol{H})$ exterior to S continue to define a solution of the field equations, hence a field, in the interior of any S_i; but since a potential of a double layer

$$f(xyz) = \iint\limits_S \mu(S)\,[d(1/r)\,dn]\; dS$$

has the property that

$$f_{n=+0} - f_{n=-0} = 4\pi\mu(S_0)$$

it results that if $\boldsymbol{n}$ continues to represent the exterior normal, instead of $4\pi\boldsymbol{H}_{u_0}$ on the left as in (1.1),

$$0 = \lim_{n=-0} \iint\limits_S (\boldsymbol{H}_u' d\phi/dn)\, dS + \iint\limits_S (B\boldsymbol{H}_u' - A\boldsymbol{H}_v')\, dS.$$

Hence the field defined in the interior of any S_i must satisfy the condition $\boldsymbol{H}_v = \boldsymbol{H}_u = 0$. The tangential components of $\boldsymbol{E}$ are the same on both sides of S_i, hence unless $[\boldsymbol{n}, \boldsymbol{E}]$ vanishes identically exterior to each S_i, on at least one surface the interior tangential components of $\boldsymbol{E}$ are not identically zero; let this surface be S_k. Interior to S_k the tangential components of $\boldsymbol{H}$ are zero, while $[\boldsymbol{n}, \boldsymbol{E}] \neq 0$. Now let the field equations be written

$$\lambda \boldsymbol{E} = \operatorname{curl} \boldsymbol{H}, \qquad \kappa \boldsymbol{H} = \operatorname{curl} \boldsymbol{E}.$$

Then

$$\kappa\lambda \boldsymbol{E} = \operatorname{curl} \kappa \boldsymbol{H}, \qquad \lambda\kappa \boldsymbol{H} = \operatorname{curl} \lambda \boldsymbol{E}$$

or if

$$\boldsymbol{E}_1 = \kappa \boldsymbol{H}, \qquad \boldsymbol{H}_1 = \lambda \boldsymbol{E},$$

the first equations are transformed into

$$\kappa \boldsymbol{H}_1 = \operatorname{curl} \boldsymbol{E}_1, \qquad \lambda \boldsymbol{E}_1 = \operatorname{curl} \boldsymbol{H}_1.$$

Hence a field $(\boldsymbol{E}_1, \boldsymbol{H}_1)$ exists, under the preceding conditions, such that interior to S_k the tangential components of $\boldsymbol{H}_1$ are not identically zero, while the tangential components of $\boldsymbol{E}_1$ vanish. This is only possible if the time-constant p is one of the set of discrete values for which cavity-radiation (Hohlraumstrahlung) exists, while the conductivity σ of the exterior region, also of the interior of S_k which is now merely a geometrical surface, is zero.

If p is not one of the set corresponding to cavity-radiation for any surface S_i, the assumption that $[\boldsymbol{n}, \boldsymbol{E}] \neq 0$ exterior to each S_i is reduced to an absurdity.

The assumption that the tangential vector $[\boldsymbol{n}, \boldsymbol{E}]$ defined by a solution of the homogeneous integral equations is identically zero on each S_i also leads to a contradiction; to show this it is necessary to derive expressions for the field at infinity.

2. From (1.4), Part I and § 3, a diverging field $(\boldsymbol{E}, \boldsymbol{H})$ can be represented in the form

$$4\pi \boldsymbol{E} = \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{E}]\, \phi dS + 1/\lambda \operatorname{curl} \operatorname{curl} \iint [\boldsymbol{n}, \boldsymbol{H}]\, \phi dS$$

$$4\pi \boldsymbol{H} = \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{H}]\, \phi dS + 1/\kappa \operatorname{curl} \operatorname{curl} \iint_S [\boldsymbol{n}, \boldsymbol{E}]\, \phi dS.$$

Let (xyz) denote a point in space, while (x_1, y_1, z_1) is a point on S. From the definition,

$$\phi = e^{-hr}/r, \qquad r = (R^2 + r_1^2 - 2Rr_1 \cos \psi)^{1/2},$$

$$r_1 = (x_1^2 + y_1^2 + z_1^2)^{1/2}, \quad R + (x^2 + y^2 + z^2)^{1/2}, \quad \cos \psi = (xx_1 + yy_1 + zz_1)/r_1R.$$

Then

$$r = R - r_1 \cos \psi + (\cdot \cdot \cdot)/R.$$

Let

$$\cos\alpha = x/R, \quad \cos\beta = y/R, \quad \cos\gamma = z/R.$$

Then the limits are easily obtained

$$\lim_{R\to\infty} Re^{hR}\phi = e^{hr_1\cos\psi},$$
$$\lim_{R\to\infty} Re^{hR}\partial\phi/\partial x = -h\cos\alpha e^{hr_1\cos\psi},$$
$$\lim_{R\to\infty} Re^{hR}\partial^2\phi/\partial x^2 = h^2\cos^2\alpha e^{hr_1\cos\psi},$$
$$\lim_{R\to\infty} Re^{hR}\partial^2\phi/\partial x\partial y = h^2\cos\alpha\cos\beta e^{hr_1\cos\psi}, \quad \text{etc.}$$

Let

$$\boldsymbol{U} = 1/4\pi \iint_S [\boldsymbol{n}, \boldsymbol{H}] e^{hr_1\cos\psi} dS, \qquad \boldsymbol{V} = 1/4\pi \iint_S [\boldsymbol{n}, \boldsymbol{E}] e^{hr_1\cos\psi} dS.$$

With the preceding limits,

$$\lim_{R\to\infty} Re^{hR}\,\mathrm{curl} \iint_S \{[\boldsymbol{n}, \boldsymbol{H}]/4\pi\}\phi dS$$
$$= -h\{i(\boldsymbol{U}_z\cos\beta - \boldsymbol{U}_y\cos\gamma) + j(\boldsymbol{U}_x\cos\gamma - \boldsymbol{U}_z\cos\alpha) + k(\boldsymbol{U}_y\cos\alpha - \boldsymbol{U}_x\cos\beta)\}.$$

Let Ω be a sphere of unit radius about the origin, while Σ is a sphere with center at the origin, and radius R. The exterior normal on each is $\bar{\boldsymbol{n}} = (\cos\alpha, \cos\beta, \cos\gamma)$; the right-hand member above is equal to

$$-h\,[\bar{\boldsymbol{n}}, \boldsymbol{U}].$$

If a vector $\boldsymbol{F}$ satisfies the equation $(\Delta - h^2)\boldsymbol{F} = 0$,

$$\mathrm{curl}\,\mathrm{curl}\,\boldsymbol{F} = \mathrm{grad}\,\mathrm{div}\,\boldsymbol{F} - h^2\boldsymbol{F};$$

consequently from the limits preceding it is found that

$$\lim_{R\to\infty} Re^{hR}\boldsymbol{E} = -h[\bar{\boldsymbol{n}}, \boldsymbol{V}] - \kappa\{\bar{\boldsymbol{n}}(\bar{\boldsymbol{n}}\boldsymbol{U}) - \boldsymbol{U}\} = \boldsymbol{E}_1$$
$$\lim_{R\to\infty} Re^{hR}\boldsymbol{H} = -h[\bar{\boldsymbol{n}}, \boldsymbol{U}] - \lambda\{\bar{\boldsymbol{n}}(\bar{\boldsymbol{n}}\boldsymbol{V}) - \boldsymbol{V}\} = \boldsymbol{H}_1.$$

Also,

$$\lim_{R\to\infty} Re^{hR}[\bar{\boldsymbol{n}}, \boldsymbol{E}] = \kappa\,[\bar{\boldsymbol{n}}, \boldsymbol{U}] + h\boldsymbol{V} - h\bar{\boldsymbol{n}}(\bar{\boldsymbol{n}}\boldsymbol{V}) = [\bar{\boldsymbol{n}}, \boldsymbol{E}_1]$$
$$\lim_{R\to\infty} Re^{hR}[\bar{\boldsymbol{n}}, \boldsymbol{H}] = \lambda\,[\bar{\boldsymbol{n}}, \boldsymbol{V}] + h\boldsymbol{U} - h\bar{\boldsymbol{n}}(\bar{\boldsymbol{n}}\boldsymbol{U}) = [\bar{\boldsymbol{n}}, \boldsymbol{H}_1]$$

and from the relation $h^2 = -\kappa\lambda$,

$$h[\bar{\boldsymbol{n}}, \boldsymbol{E}_1] = -\kappa\boldsymbol{H}_1, \qquad h[\bar{\boldsymbol{n}}, \boldsymbol{H}_1] = -\lambda\boldsymbol{E}_1.$$

Let the conductivity of the exterior region be zero, while $p = i\omega$; then $\lambda = \epsilon i\omega/c$, $\kappa = -\mu i\omega/c$, $h = i\omega(\epsilon\mu)^{1/2}/c$. The preceding identities become

$$(\epsilon\mu)^{1/2}[\bar{\boldsymbol{n}}, \boldsymbol{E}_1] = \mu \boldsymbol{H}_1, \qquad (\epsilon\mu)^{1/2}[\bar{\boldsymbol{n}}, \boldsymbol{H}_1] = -\epsilon \boldsymbol{E}_1.$$

Now the energy flow outward across the sphere Σ is equal to

$$c/4\pi \iint_{\Sigma} \{R[\boldsymbol{E}e^{i\omega t}], \quad R[\boldsymbol{H}e^{i\omega t}]\}_{\bar{n}}\, dS$$

if only the real parts of the field are considered; omitting the factor $c/4\pi$, the average value of this expression over a period $T = 2\pi/\omega$ is half the quantity

$$\begin{aligned} K_R &= \tfrac{1}{2} \iint_{\Sigma} \{[\boldsymbol{E}, \bar{\boldsymbol{H}}]_{\bar{n}} + [\bar{\boldsymbol{E}}, \boldsymbol{H}]_{\bar{n}}\} dS \\ &= -\tfrac{1}{2} \iint_{\Sigma} \{\boldsymbol{E}[\bar{\boldsymbol{n}}, \bar{\boldsymbol{H}}] + \bar{\boldsymbol{E}}[\bar{\boldsymbol{n}}, \boldsymbol{H}]\} dS \\ &= -\tfrac{1}{2} \iint_{\Omega} \{Re^{hR}\boldsymbol{E}[\bar{\boldsymbol{n}}, Re^{-hR}\bar{\boldsymbol{H}}] + Re^{-hR}\bar{\boldsymbol{E}}\,[\bar{\boldsymbol{n}}, Re^{hR}\boldsymbol{H}]\} d\omega. \end{aligned}$$

From the identities developed above it is seen that

$$\begin{aligned} \lim_{R\to\infty} K_R &= -\tfrac{1}{2} \iint_{\Omega} \{\boldsymbol{E}_1[\bar{\boldsymbol{n}}, \boldsymbol{H}_1] + \boldsymbol{E}_1[\bar{\boldsymbol{n}}, \boldsymbol{H}_1]\} d\omega \\ &= (\epsilon/\mu)^{1/2} \iint_{\Omega} |\, \boldsymbol{E}_1 \,|^2\, d\omega = (\mu/\epsilon)^{1/2} \iint_{\Omega} |\, \boldsymbol{H}_1 \,|^2\, d\omega. \end{aligned}$$

It is an immediate consequence that if the mean energy flow at infinity is zero, the functions $\boldsymbol{E}_1$, $\boldsymbol{H}_1$ vanish identically. This condition is satisfied if the tangential components of $\boldsymbol{E}$ or of $\boldsymbol{H}$ are zero exterior to the system of surfaces S; for in the identity *

$$\iiint_{(S+\Sigma)} \{\operatorname{curl} \boldsymbol{u} \operatorname{curl} \boldsymbol{v} + \operatorname{div} \boldsymbol{u} \operatorname{div} \boldsymbol{v} + \boldsymbol{u}\Delta\boldsymbol{v}\} dx\, dy\, dz$$
$$= -\iint_{(S+\Sigma)} \{[\boldsymbol{n}, \boldsymbol{u}] \operatorname{curl} \boldsymbol{v} + \boldsymbol{u}_n \operatorname{div} \boldsymbol{v}\} dS$$

replace $\boldsymbol{u}$ by $\boldsymbol{H}$, $\boldsymbol{v}$ by $\bar{\boldsymbol{H}}$; from the result subtract the corresponding members of the identity with $\boldsymbol{u} = \bar{\boldsymbol{H}}$, $\boldsymbol{v} = \boldsymbol{H}$. Since

$$\operatorname{curl} \boldsymbol{H} = -\lambda \bar{\boldsymbol{E}},$$

the right-hand member becomes

$$\lambda \iint \{\bar{\boldsymbol{E}}[\boldsymbol{n}, \boldsymbol{H}] + \boldsymbol{E}[\boldsymbol{n}, \bar{\boldsymbol{H}}]\} dS$$

and vanishes since the left-hand member is zero. Consequently if the inte-

* Weyl, *loc. cit.*

grand of this expression, which depends only on the tangential components of $\boldsymbol{E}, \bar{\boldsymbol{E}}, \boldsymbol{H}, \bar{\boldsymbol{H}}$ vanishes on S, the integral is also zero on Σ, for every $R > R_0$.

If $\boldsymbol{E}_1 = \boldsymbol{H}_1 = 0$, it results that the field $(\boldsymbol{E}, \boldsymbol{H})$ vanishes identically exterior to S. For the components of the field satisfy the wave equation, which becomes in spherical coördinates

$$\frac{1}{R^2}\,\frac{\partial}{\partial R}\left(R^2\frac{\partial u}{\partial R}\right) + \frac{1}{R^2 \sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) + \frac{1}{R^2\sin^2\theta}\,\frac{\partial^2 u}{\partial\phi^2} - h^2 u = 0.$$

Let Y_n be an arbitrary spherical harmonic of order n; the integral

$$W_n(R) = \int_0^{2\pi} d\phi \int_0^{\pi} \boldsymbol{E}_x(R,\theta,\phi)\,Y_n(\theta,\phi)\sin\theta d\theta$$

satisfies the differential equation

$$\frac{1}{R^2}\,\frac{d}{dR}\left(R^2\frac{dW_n}{dR}\right) + \left(-h^2 - \frac{n(n+1)}{R^2}\right)W_n = 0$$

as is seen from the equation in Y_n,

$$\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y_n}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\,\frac{\partial^2 Y_n}{\partial\phi^2} + n(n+1)Y_n = 0$$

and an integration by parts.*

Consequently

$$\begin{aligned} W_n &= C_1 K_{n+1/2}\ (hR)/(R)^{1/2} + C_2 I_{n+1/2}\ (hR)/(R)^{1/2} \\ &= C_1/(R)^{1/2}\,(\pi/2hR)^{1/2}\,e^{-hR}\,[1 + 1/R(\cdot\ \cdot\ \cdot)] \end{aligned}$$

since a limit must exist, as $R \to \infty$, for the quantity $Re^{hR}W_n$.

$$\lim_{R\to\infty} Re^{hR}W_n = C_1{}^{(n)}(\pi/2h)^{1/2}.$$

Since the products $Re^{hR}\boldsymbol{E}_x$, etc., approach their limits uniformly with respect to the spherical angles θ, ϕ,

$$\begin{aligned} &\lim_{R\to\infty} Re^{hR} \iint_\Omega \boldsymbol{E}_x(R,\theta,\phi)Y_n(\theta,\phi)\sin\theta\,d\theta d\phi \\ &\qquad = \iint_\Omega Y_n(\theta,\phi)\boldsymbol{E}_{1x}(\theta,\phi)\sin\theta\,d\theta d\phi = C_1{}^{(n)}(\pi/2h)^{1/2}. \end{aligned}$$

If $\boldsymbol{E}_1 = 0$, $C_1 = 0$. But Y_n was an arbitrary spherical harmonic, hence the expansions of the components of $\boldsymbol{E}$ in spherical harmonics on the surface Σ vanish identically. Since $\boldsymbol{E}$ is analytic on Σ, provided R is so large that the surfaces S_i lie entirely in the interior of Σ, $\boldsymbol{E}$ must vanish identically

* Carleman, *Sur les équations intégrales singulières á noyau réel et symetrique*, Upsala, 1923, pp. 181-183.

on Σ, likewise $\boldsymbol{H}$. If the surfaces S_i are ordinary surfaces, the values of $\boldsymbol{E}$ and $\boldsymbol{H}$ at an ordinary point exterior to S can be obtained by analytic continuation from their values on some Σ, hence the field $(\boldsymbol{E}, \boldsymbol{H})$ vanishes identically exterior to S.

Returning to the homogeneous equations assumed satisfied when the surfaces S_i bound perfect conductors, it has been found that the field $(\boldsymbol{E}, \boldsymbol{H})$ defined by a solution of these equations must vanish identically exterior to S, if p is not one of the exceptional values. But if $\boldsymbol{H}$ vanishes identically exterior to S, the tangential components of $\boldsymbol{H}$ on S vanish identically; the integral equations express that the vector $[\boldsymbol{n}, \boldsymbol{H}]$ appearing in the representation of $\boldsymbol{H}$ is formed from these tangential components, hence vanishes identically. The assumption that a solution of the homogeneous equations exists leads therefore to a contradiction unless p is one of the exceptional values corresponding to cavity-radiation in the interior of some S_i.

From the theorems of Fredholm it results that the equations for the tangential components of $(\boldsymbol{E}, \boldsymbol{H})$ when an impressed field is present have a unique solution, if p is not one of these exceptional values.

THE GRAVITATIONAL FIELD OF A BODY WITH ROTATIONAL SYMMETRY IN EINSTEIN'S THEORY OF GRAVITATION.

By P. Y. Chou.

Introduction.

The present paper is an attempt to solve rigorously the problem of the static gravitational field of a body whose mass is distributed symmetrically around an axis in Einstein's theory of gravitation. In § 1 Einstein's field equations in vacuo *

$$G_{\mu\nu} = 0 \tag{0.1}$$

are set up and reduced in § 2 to a form such that simple problems like the sphere (§ 4) and the plane (§ 5) can be solved. In the general problem there is a fundamental difficulty which will be avoided by the introduction of the Newtonian potential (§ 7). The solution of the whole problem then depends upon the solution of the well known Laplace's equation and a partial differential equation of the second order which is not linear. Finally the gravitational fields of spheroidal homoeoids (§ 8, § 9) are given as illustrations of the present investigation and the motion of a particle in the field of an oblate spheroidal homoeoid is discussed (§ 10). The paper also contains a critical examination of earlier works upon the problem notably those of Prof.'s Weyl and Levi-Civita (§ 3).

I. Einstein's Law of Gravitation.

1. *The field equations.* We consider the static gravitational field outside of a body whose mass is distributed symmetrically about an axis. Hence the $g_{\mu\nu}$'s do not vary with respect to time. The most general fundamental quadratic differential form in such a field appears to be

$$ds^2 = -(g_{11}dx_1^2 + 2g_{12}dx_1dx_2 + g_{22}dx_2^2) - g_{33}dx_3^2 + g_{44}dx_4^2 \tag{1.1}$$

where x_1, x_2 are any two coördinates in the meridianal plane containing the z-axis, $x_3 = \phi$, the azimuthal angle, $x_4 = t$, the time coördinate, the unit of time being so chosen that the velocity of light in vacuo is unity. The $g_{\mu\nu}$'s in (1.1) are functions of x_1 and x_2 only.

* A. S. Eddington, *The Mathematical Theory of Relativity*, 2nd Ed. (1924), p. 81. Eddington's notation with slight modifications will be followed throughout the present paper.

We assume that the values of the $g_{\mu\nu}$'s exist. From a well-known theorem * on positive definite quadratic differential forms of two variables such as the form in the parenthesis of (1.1), it is always possible when g_{11}, g_{12}, g_{22} are explicitly given, to make a real single-valued, continuous transformation from x_1 and x_2 to u and v by

$$(1.2)\qquad x_1 = x_1(u,v),\quad x_2 = x_2(u,v),\quad \text{where } J \equiv [\partial(x_1,x_2)/\partial(u,v)] \neq 0$$

such that the following identity is true,

$$(1.3)\qquad g_{11}dx_1^2 + 2g_{12}dx_1dx_2 + g_{22}dx_2^2 \equiv e^{2m}(du^2 + dv^2).$$

Hence (1.1) becomes

$$(1.4)\qquad ds^2 = -e^{2m}(du^2 + dv^2) - e^{2n}dx_3^2 + e^{2\nu}dx_4^2$$

where m, n, ν are functions of u and v to be determined. Let

$$u = x_1, \quad v = x_2.$$

Then

$$(1.5)\qquad \begin{aligned} g_{11} = g_{22} &= -e^{2m}, \quad g_{33} = -e^{2n}, \quad g_{44} = e^{2\nu}, \\ g &= g_{11}g_{22}g_{33}g_{44} = -e^{4m+2n+2\nu} \\ g^{11} = g^{22} &= -e^{-2m}, \quad g^{33} = -e^{-2n}, \quad g^{44} = e^{-2\nu}. \end{aligned}$$

Now (1.4) is an orthogonal quadratic differential form. The general expressions of the Christoffel symbols of the second kind for such forms are well known.† In the present problem the non-vanishing symbols are

$$(1.6)\qquad \begin{aligned} &\{11,1\} = m_u && \{11,2\} = -m_v \\ &\{12,1\} = m_v && \{12,2\} = m_u \\ &\{22,1\} = -m_u && \{22,2\} = m_v \\ &\{33,1\} = -e^{2n-2m}n_u && \{33,2\} = -e^{2n-2m}n_v \\ &\{44,1\} = e^{2\nu-2m}\nu_u && \{44,2\} = e^{2\nu-2m}\nu_v \\ &\{13,3\} = n_u && \{14,4\} = \nu_u \\ &\{23,3\} = n_v && \{24,4\} = \nu_v \end{aligned}$$

where the subscripts mean partial differentiations for simplicity.

Written out in full Einstein's field equations in vacuo are

$$(1.7)\qquad \begin{aligned} G_{\mu\nu} \equiv &-\partial/\partial x_a\{\mu\nu,\alpha\} + \{\mu\alpha,\beta\}\{\nu\beta,\alpha\} \\ &\partial^2/\partial x_\mu\partial x_\nu \log(-g)^{1/2} - \{\mu\nu,\alpha\}\,\partial/\partial x_a \log(-g)^{1/2} = 0. \end{aligned}$$

If we substitute for the three-index Christoffel symbols of the second kind

---

* L. Bianchi-Lucat, *Vorlesungen über Differentialgeometrie*, (1910), pp. 69.

† A. S. Eddington, *loc. cit.*, pp. 83.

from their values (1. 6) into (1. 7), we obtain the following five non-vanishing components,

$$(1.8)\quad G_{11} \equiv m_{uu} + m_{vv} + n_{uu} + \nu_{uu} + n_u^2 + \nu_u^2 - m_u(n_u + \nu_u) + m_v(n_v + \nu_v) = 0,$$
$$(1.9)\quad G_{12} \equiv n_{uv} + \nu_{uv} + n_u n_v + \nu_u \nu_v - m_v(n_u + \nu_u) - m_u(n_v + \nu_v) = 0,$$
$$(1.10)\quad G_{22} \equiv m_{uu} + m_{vv} + n_{vv} + \nu_{vv} + n_v^2 + \nu_v^2 + m_u(n_u + \nu_u) - m_v(n_v + \nu_v) = 0,$$
$$(1.11)\quad G_{33} \equiv e^{2n-2m}\,[n_{uu} + n_{vv} + n_u(n_u + \nu_u) + n_v(n_v + \nu_v)] = 0,$$
$$(1.12)\quad G_{44} \equiv -e^{2\nu-2m}\,[\nu_{uu} + \nu_{vv} + \nu_u(n_u + \nu_u) + \nu_v(n_v + \nu_v)] = 0.$$

2. *Reduction of the field equations.* By putting

$$(2.1)\qquad \chi = n + \nu,$$

and adding the expressions in the square brackets of (1. 11) and (1. 12) we get

$$(2.2)\qquad \chi_{uu} + \chi_{vv} + \chi_u^2 + \chi_v^2 = 0,$$

which becomes Laplace's equation in the uv-plane,

$$(2.3)\qquad \Phi_{uu} + \Phi_{vv} = 0, \quad \text{on setting} \quad \Phi = e^{\chi} = e^{n+\nu}.$$

It is well known that the solution of (2. 3) is unique, if the boundary value of Φ be given in the uv-plane. Then $G_{44} = 0$ becomes

$$(2.4)\qquad \nu_{uu} + \nu_{vv} + \chi_u \nu_u + \chi_v \nu_v = 0,$$

which determines ν. We obtain n by (2. 1).

To get the unknown function, m, we use (1. 8), (1. 9), (1. 10). Write

$$(2.5)\qquad \begin{aligned} &G_{12} = 0: \\ &\qquad \chi_v m_u + \chi_u m_v = \chi_{uv} + n_u n_v + \nu_u \nu_v \equiv A, \\ &G_{11} - G_{22} = 0: \\ &\qquad -\chi_u m_u + \chi_v m_v = \tfrac{1}{2}\{-\chi_{uu} + \chi_{vv} - n_u^2 - \nu_u^2 + n_v^2 + \nu_v^2\} \equiv B. \end{aligned}$$

Then by solving m_u and m_v simultaneously from (2. 5) we get

$$(2.6)\qquad m_u = (\chi_u^2 + \chi_v^2)^{-1}\{\chi_v A - \chi_u B\}, \quad m_v = (\chi_u^2 + \chi_v^2)^{-1}\{\chi_u A + \chi_v B\}.$$

It can be shown by direct differentiation and the aid of (1. 11) and (1. 12) that

$$(2.7)\qquad dm \equiv m_u du + m_v dv$$

where m_u and m_v are given in (2. 6) is an exact differential and secondly that m satisfy (1. 8) and (1. 10). This completes the proof that the functions, m, n, ν, thus obtained satisfy every component of Einstein's field equations (1. 7).

3. *On Weyl-Levi-Civita's solution.* The problem under consideration was first attacked by H. Weyl.* His result was criticized by T. Levi-Civita † as being incomplete due to the incomplete use of a variational principle. The latter started with (0. 1) and gave a complete though restricted set of solutions. Consider (2. 3) and set $\Phi = \rho$. Let z be the conjugate function of ρ. Then

$$\rho + iz = f(u + iv) \tag{3.1}$$

where $f(u + iv)$ is analytic in $u + iv$. From this it follows that

$$d\rho^2 + dz^2 = f'(u + iv)f'(u - iv)(du^2 + dv^2) \tag{3.2}$$

namely, one set of coördinates (say u, v) is conformally transformed into the other (say ρ, z). In order to avoid cumbersome mathematical manipulations in § 2, both Weyl and Levi-Civita assume initially that

$$\rho = u, \quad z = v. \quad \text{Then} \quad e^{2n} = \rho^2 e^{-2\nu} \quad \text{and} \tag{3.3}$$

$$ds^2 = -e^{2m}(d\rho^2 + dz^2) - \rho^2 e^{-2\nu} d\phi^2 + e^{2\nu} dt^2. \tag{3.4}$$

Moreover, $G_{44} = 0$ and dm become respectively

$$\partial^2\nu/\partial\rho^2 + \partial^2\nu/\partial z^2 + (1/\rho)\partial\nu/\partial\rho = 0, \tag{3.5}$$

$$dm = -d\nu + \rho[(\partial\nu/\partial\rho)^2 - (\partial\nu/\partial z)^2]\, d\rho + 2\rho\, (\partial\nu/\partial\rho)(\partial\nu/\partial z) dz. \tag{3.6}$$

We recognize (3. 5) as Laplace's equation in cylindrical coördinates (ρ, z, ϕ) independent of ϕ. Weyl calls (ρ, z) in (3. 3) the "canonical cylindrical coördinates" which are apparently different from the ordinary cylindrical coördinates used in solving Newtonian potential problems. He then emphasizes the fact ‡ that if the distribution of mass of a given body in our space-time manifold is known in terms of this set of configurational canonical coördinates, the problem is reduced to the solution of (3. 5). He shows § that Schwarzschild's solution in isotropic coördinates of a body with mass, m, having spherical symmetry, corresponds to that of a finite line segment of length $2m$, with constant linear density, lying on the z-axis of the configurational canonical space-time manifold. But he does not make clear that it is almost impossible to know the corresponding distribution of mass in this

---

* H. Weyl, *Annalen der Physik*, Bd. 54 (1918), pp, 134.

† T. Levi-Civita, *Rendiconti Accademie dei Lincei*, Vol. 28, i (1919), pp. 9.

‡ H. Weyl, *Raum, Zeit, Materie*, 5th ed. (1923), p. 266.

§ H. Weyl, *Annalen der Physik*, Bd. 54 (1918), p. 140. Schwarzschild's solution is not necessarily limited to a particle. It can be applied equally well to a spherical shell. Cf. J. T. Combridge, *Philosophical Magazine* (7), Vol. 1 (1926), pp. 276.

canonical coördinate system when the distribution of mass in our space-time coördinates is given. This difficulty is clearly brought out by the following argument.

When we carry out the transformation from (x_1, x_2, ϕ, t) to (ρ, z, ϕ, t) by (1. 2) and (3. 3) we assume only the existence of the values of g_{11}, g_{12}, g_{22} in (1. 1) so that the transformation is possible, but their explicit forms are not given *a priori* and consequently (1. 2) is not explicitly known. Although we know the boundary values of $g_{\mu\nu}$ in the original (x_1, x_2, ϕ, t) system, we do not know the corresponding boundary conditions in the (u, v, ϕ, t) system on account of the uncertainty of (1. 2). Since (3. 5) has an infinite number of solutions if the boundary value of ν is not specified, the solution obtainable from (3. 5) and (3. 6) will not be unique, and consequently it is indeterminate.

The same difficulty arises if we do not assume the solution of Φ in (3. 3). Here we do not know which solution of (2. 3) we should take in order to solve (2. 4). The complexity of the situation is enhanced further by the uncertainty of the boundary conditions of ν in the uv-coördinates.

An alternative procedure to get a solution for the original physical problem from (3. 5) and (3. 6) is to choose a solution of (3. 5) in terms of the canonical coördinates first and then try to interpret it in the (x_1, x_2, ϕ, t) system by a transformation (1. 2). The $g_{\mu\nu}$'s thus obtained must satisfy the original boundary conditions in terms of (x_1, x_2, ϕ, t) given initially. The question whether this procedure will lead to a unique transformation (1. 2) needs further investigation. It appears not to have been considered in the literature.

II. Fields of Sphere and Plane.

4. *Schwarzschild's solution.* As the first application of the results in § 2 let us consider Schwarzschild's solution. The arc element in the gravitational field outside a body with spherical symmetry is

$$(4.1)\qquad ds^2 = -e^{2\lambda}dr^2 - e^{2\mu}(r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2) + e^{2\nu}dt^2,$$

where (r, θ, ϕ) denote spherical polar coördinates and λ, μ, ν are functions of r only. (4. 1) may be put in the form of (1. 4),

$$(4.2)\qquad ds^2 = -e^{2m}(du^2 + dv^2) - e^{2n}d\phi^2 + e^{2\nu}dt^2, \quad \text{where}$$

$$(4.3)\qquad du = r^{-1}e^{\lambda-\mu}dr, \quad v = \theta, \quad e^m = re^\mu, \quad e^n = r \sin\theta e^\mu,$$

m, λ, ν being then functions of u. Let $\Phi = e^{n+\nu} \equiv R \sin v$. Then (2. 3) is

$$(4.4)\qquad d^2R/du^2 - R = 0.$$

Integrating (4.4), we obtain

(4.5) $$R = re^{\mu+\nu} = c_1 \sinh u.$$

Then (2.4) becomes $d^2\nu/du^2 + \coth u(d\nu/du) = 0$ giving

(4.6) $$\exp(\nu/c_2) = c_3(\coth u - \operatorname{csch} u).$$

Eliminating u between (4.5) and (4.6) we obtain

(4.7) $$\exp(\nu/c_2) = c_3 r^{-1} e^{-(\mu+\nu)}\{[c_1^2 + r^2 e^{2(\mu+\nu)}]^{1/2} - c_1\}.$$

By using the boundary condition on μ and ν that both of them tend toward zero as r increases indefinitely we get $c_3 = 1$. Solving ν from (4.7), we get

(4.8) $$\sinh(\nu/c_2) = -c_1 r^{-1} e^{-(\mu+\nu)} = -\operatorname{csch} u.$$

From (4.3) we see that m is a function of u only. In (2.6) we must have $m_v = 0$ which together with (4.6) gives

(4.9) $$c_2 = \pm 1.$$

Take $c_2 = 1$. Then from (4.8) we have

(4.10) $$e^{2\nu} = 1 - (2c_1/r)e^{-\mu}.$$

Eliminating u between (4.3) and (4.6), we obtain

(4.11) $$e^{\lambda} = -r \operatorname{csch} \nu (d\nu/dr) e^{\mu}.$$

The second case $c_2 = -1$ only changes c_1 to $-c_1$.

Relations (4.10) and (4.11) connect the three unknown functions λ, μ, ν. Consequently an infinite number of solutions arises. To obtain Schwarzschild's solution we set $\mu = 0$. Then (4.10) becomes

(4.12) $$g_{44} = e^{2\nu} = 1 - 2c_1/r$$

where c_1 may be identified as the mass of the body from Newton's theory. From (4.10) and (4.11) it follows that $\lambda = -\nu$. The same result can be also obtained by assuming that g_{44} is $1 - 2V$ to start with where V is the Newtonian potential of the body.

A second solution of interest is the one in isotropic coördinates where the velocity of light is independent of direction. Putting $\lambda = \mu$ in (4.11) and integrating, we get

(4.13) $$\sinh \nu = 2c_4 r(r^2 - c_4^2)^{-1}.$$

To determine the constant of integration, c_4, we use (4.10) and let r tend toward infinity. This gives

(4.14) $$2c_4 = -c_1.$$

Solving for e^{ν} and rejecting the negative root of e^{ν} which is essentially positive from (4.13), we find

(4.15) $$g_{44} \equiv e^{2\nu} = (2r - c_1)^2/(2r + c_1)^2 \quad \text{and} \quad e^{2\mu} = (1 + c_1/2r)^4.$$

This result was also obtained by a transformation of r in Schwarzschild's solution.*

5. *Infinite plane: a. Gravitational field.* Let the xy-plane be the given plane. From symmetry considerations around any line parallel to the z-axis the most general fundamental quadratic differential form appears to be

(5.1) $$ds^2 = -e^{2\lambda}(d\rho^2 + \rho^2 d\phi^2) - e^{2\mu}dz^2 + e^{2\nu}dt^2$$

where λ, μ, ν are functions of z only. (5.1) can be put in the form (1.4),

(5.2) $$ds^2 = -e^{2m}(du^2 + dv^2) - e^{2n}d\phi^2 + e^{2\nu}dt^2 \quad \text{with}$$

(5.3) $$du = e^{\mu-\lambda}dz, \quad \rho = v, \quad m = \lambda, \quad e^{2n} = \rho e^{\lambda}.$$

In the present case $\Phi = \rho e^{\lambda+\nu} \equiv \rho R$ and (2.3) becomes

(5.4) $$d^2R/du^2 = 0. \quad \text{Hence}$$

(5.5) $$R = e^{\lambda+\nu} = c_1 u.$$

Then (2.4) becomes $d^2\nu/du^2 + (1/u)d\nu/du = 0$ giving

(5.6) $$\nu = \log c_3 + c_2 \log u.$$

From (5.3) and $n = \chi - \nu$, we find

(5.7) $$\lambda = (1 - c_2)\log u + \log c_4, \qquad (c_4 = c_1/c_3).$$

By (2.6), $m_v = \lambda_v = 0$, we get

(5.8) $$c_2 = \pm 1.$$

Consider $c_2 = 1$. If we choose the unit of length properly, $c_3 = c_1$ and $e^{\lambda} = 1$. Then (5.5) becomes

(5.9) $$e^{\nu} = c_1 u.$$

Differentiating (5.9) and on using (5.3) we find

(5.10) $$e^{\nu}(d\nu/dz) = c_1 e^{\mu}.$$

---

* A. S. Eddington, *loc. cit.*, p. 93.

Here we have again an infinite number of solutions of ν and μ. To avoid this indeterminateness we use Newton's theory. By setting

$$g_{44} \equiv e^{2\nu} = 1 + 4\pi\sigma z \tag{5.11}$$

and identifying c_1 as $2\pi\sigma$ where σ is the surface density of matter on the given plane, we get $\mu + \nu = 0$ and the final form of (5.1) is

$$ds^2 = -(1 + 4\pi\sigma z)^{-1}dz^2 - (d\rho^2 + \rho^2 d\phi^2) + (1 + 4\pi\sigma z)\,dt^2. \tag{5.12}$$

The additive constant in (5.11) is chosen to be unity. Here we are dealing with a body whose mass extends to an infinite distance and ds^2 is not Galilean at infinity. The latter condition, however, can be replaced by the one that space surrounding the plane is flat if the density of matter on the plane vanishes. This is satisfied by (5.12).

The solution (5.12) can be regarded as the limiting case of Schwarzschild's solution of a spherical shell when the radius of the shell becomes infinitely great (neglecting the infinite constant obtained in this limiting process). In fact Whittaker * uses this method to obtain his "quasi-uniform" gravitational field which, as we see in the present discussion, is the field outside an infinite material plane. The case $c_2 = -1$ and hence $e^\lambda = c_4 u^2$ has been treated by Levi-Civita † and the result extended to the gravitational field of a charged plane by Kar.‡

b. *The motion of a particle in the field.* Now it is possible to transform away the field represented by (5.12). Let

$$\begin{aligned} x' &= x, \quad y' = y, \\ z' &= (1/2\pi\sigma)\{(1 + 4\pi\sigma z)^{1/2}\cos h\, 2\pi\sigma t - (1 + 4\pi\sigma h)^{1/2}\}, \\ t' &= (1/2\pi\sigma)(1 + 4\pi\sigma z)^{1/2}\sin h\, 2\pi\sigma t, \end{aligned} \tag{5.13}$$

where h is a constant and when $(x, y, z, t) = (0, 0, h, 0)$, we have $(x', y', z', t') = (0, 0, 0, 0)$. Then (5.12) becomes

$$ds^2 = -dx'^2 - dy'^2 - dz'^2 + dt'^2. \tag{5.14}$$

If we expand (5.13) in terms of σ, we get

$$\begin{aligned} x' &= x, \quad y' = y, \\ z' &= z - h + \tfrac{1}{2}gt^2 + (g/2c^2)(h^2 - z^2 + gzt^2) + \cdots, \\ t' &= t + (g/c^2)(zt + \tfrac{1}{6}gt^3) + \cdots. \end{aligned} \tag{5.15}$$

---

* E. T. Whittaker, *Proceedings of the Royal Society* (A), Vol. 116 (1927), p. 722.
† T. Levi-Civita, *Accademie dei Lincei*, Vol. 27, ii (1918), pp. 240.
‡ S. C. Kar, *Physik Zeit*, Vol. 27 (1926), pp. 208.

where c, the velocity of light in vacuo is restored and $g = 2\pi\sigma$, the gravity. From this we see that (x, y, z, t) are the coördinates used by an observer on the plane and (x', y', z', t') those of one who falls freely toward the plane at an initial height h. In the (x', y', z', t') system gravitation of the plane vanishes.

From (5.13) and (5.14) it is obvious that the differential equations defining the geodesics in the field (5.12) can be integrated rigorously. Analysis shows that the orbit of an infinitesimal particle in the (x, y, z, t) system is a parabola, though in the primed system it is a straight line. Furthermore, this Einsteinian parabola is slightly different from that predicted according to Newton's theory, but when the velocity of light in vacuo becomes infinite, this difference disappears.

III. General Solution of the Problem.

6. *Transformation of the fundamental quadratic differential form.* The foregoing two special cases are solvable from (2.3), (2.4) and (2.6). This is because (2.3) degenerates into an ordinary differential equation in each case. In reality when Φ is a general function of u and v, the problem can be hardly solvable on account of the uncertainty of the boundary conditions of Φ in the (u, v, ϕ, t) manifold as we have pointed out in § 3. In the following section we shall avoid this difficulty by introducing the Newtonian potential into the present problem. As we shall see presently, the problem of the general static gravitational field of a finite body with rotational symmetry can be solved provided we can solve a non-linear partial differential equation of the second order.

We start with the cylindrical coördinates (ρ, z, ϕ), the z-axis being the axis of symmetry of the given body which is finite in extent. Consider the meridianal plane containing the z-axis. Choose in this plane as in ordinary potential theory a more general set (ξ, η) which is conformally mapped upon (ρ, z) by

$$(6.1) \qquad z + i\rho = F(\xi + i\eta)$$

where $F(\xi + i\eta)$ is a monogenic function of $\xi + i\eta$ so that

$$(6.2) \qquad dz^2 + d\rho^2 = h^2(d\xi^2 + d\eta^2), \quad h^2 \equiv F'(\xi + i\eta)F'(\xi - i\eta).$$

Let $\psi(\xi, \eta) = \text{const.}$, $\theta(\xi, \eta) = \text{const.}$, be two orthogonal (in the Euclidean sense) families of curves to be determined in the plane. Denote partial differentiations by subscripts as in § 1. Then

(6.3) $$d\psi = \psi_\xi d\xi + \psi_\eta d\eta, \quad d\theta = \theta_\xi d\xi + \theta_\eta d\eta \quad \text{where}$$

(6.4) $$\psi_\xi \theta_\xi + \psi_\eta \theta_\eta = 0.$$

Choose the Jacobian of transformation of (6.3) to be

(6.5) $$J \equiv \frac{\partial(\psi, \theta)}{\partial(\xi, \eta)} \equiv \psi_\xi \theta_\eta - \psi_\eta \theta_\xi = e^f(\psi_\xi^2 + \psi_\eta^2) \quad \text{where}$$

(6.6) $$e^f = \rho.$$

Solving θ_ξ, θ_η from (6.4) and (6.5) simultaneously we obtain

(6.7) $$\theta_\xi = - e^f \psi_\eta, \quad \theta_\eta = e^f \psi_\xi.$$

Since $d\theta$ is an exact differential, (6.7) must satisfy the necessary and sufficient condition,

(6.8) $$\frac{\partial}{\partial \xi}\theta_\eta = \frac{\partial}{\partial \eta}\theta_\xi \quad \text{giving}$$

(6.9) $$\psi_{\xi\xi} + \psi_{\eta\eta} + f_\xi \psi_\xi + f_\eta \psi_\eta = 0.$$

By (6.1) and (6.6), f is a known function of ξ and η. Simple verification shows that (6.9) is Laplace's equation in the (ξ, η, ϕ) coördinates independent of ϕ.

Now by (6.3) we obtain (6.2) in the form

(6.10) $$dz^2 + d\rho^2 = h^2(\psi_\xi^2 + \psi_\eta^2)^{-1}(d\psi^2 + \rho^{-2}d\theta^2).$$

Consequently the fundamental quadratic form for a flat space-time continuum in the present (ψ, θ, ϕ, t) variables is

(6.11) $$ds^2 = - h^2(\psi_\xi^2 + \psi_\eta^2)^{-1}(d\psi^2 + \rho^{-2}d\theta^2) - \rho^2 d\phi^2 + dt^2.$$

When matter is present, ds^2 is no more Galilean. We suppose that in such cases (6.11) is replaced by

(6.12) $$ds^2 = - e^{-2H}(e^{2\lambda}d\psi^2 + \rho^{-2}e^{2\mu}d\theta^2) - \rho^2 e^{2\gamma}d\phi^2 + e^{2\nu}dt^2,$$

(6.13) $$e^{-2H} = h^2(\psi_\xi^2 + \psi_\eta^2)^{-1},$$

where $\lambda, \mu, \gamma, \nu$ are functions of ψ and θ to be determined according to Einstein's law of gravitation, with the condition that at infinite distances from the body all four approach zero as a limit.

7. *Introduction of the Newtonian potential.* Next transform (2.3), (2.4) and (2.6) into the (ψ, θ, ϕ, t) system. Consider the following expression from (6.12),

(7.1) $$d\psi^2 + \rho^{-2}e^{-2\lambda+2\mu}d\theta^2 \equiv d\psi^2 + e^{-2g}d\theta^2 = (d\psi + ie^{-g}d\theta)(d\psi - ie^{-g}d\theta)$$

where we put $e^g = \rho e^{\lambda-\mu}$. Let $(\alpha + i\beta)^{-1} \neq 0$, where both α and β are real, be an integrating factor of $d\psi + ie^{-g}d\theta$ so that

(7.2) $\quad d\psi + ie^{-g}d\theta = (\alpha + i\beta)(du + idv)$, and (6.2) becomes

(7.3) $$ds^2 = -e^{-2H}e^{2\lambda}(\alpha^2 + \beta^2)(du^2 + dv^2) - \rho^2 e^{2\gamma}d\phi^2 + e^{2\nu}dt^2.$$

Comparing (7.3) and (1.4) we obtain

(7.4) $$e^{2m} = e^{-2H}e^{2\lambda}(\alpha^2 + \beta^2).$$

Equating real and imaginary parts in (7.2) we get

(7.5) $$d\psi = \alpha du - \beta dv, \quad d\theta = e^g(\beta du + \alpha dv),$$

from which the conditions of integrability for $d\psi$ and $d\theta$ give

(7.6) $\quad \alpha_v + \beta_u = 0, \quad \alpha_u - \beta_v = \beta g_v - \alpha g_u,$ and furthermore

(7.7) $$\psi_{uu} + \psi_{vv} = -(\alpha^2 + \beta^2)\frac{\partial g}{\partial \psi}, \quad \theta_{uu} + \theta_{vv} = (\alpha^2 + \beta^2)e^{2g}\frac{\partial g}{\partial \theta}.$$

By (7.5) and (7.7), equation (2.3) becomes

(7.8) $$\frac{\partial^2 \Phi}{\partial \psi^2} + e^{2g}\frac{\partial^2 \Phi}{\partial \theta^2} - \frac{\partial g}{\partial \psi}\,\frac{\partial \Phi}{\partial \psi} + e^{2g}\frac{\partial g}{\partial \theta}\,\frac{\partial \Phi}{\partial \theta} = 0.$$

In the like manner we get (2.4) in the form,

(7.9) $$\frac{\partial^2 \nu}{\partial \psi^2} + e^{2g}\frac{\partial^2 \nu}{\partial \theta^2} - \frac{\partial g}{\partial \psi}\,\frac{\partial \nu}{\partial \psi} + e^{2g}\frac{\partial g}{\partial \theta}\,\frac{\partial \nu}{\partial \theta} + \frac{\partial \nu}{\partial \psi}\,\frac{\partial \chi}{\partial \psi} + e^{2g}\frac{\partial \nu}{\partial \theta}\,\frac{\partial \chi}{\partial \theta} = 0.$$

In (6.12) we have four functions λ, μ, γ, ν to determine by means of the three independent equations (7.8), (7.9) and (2.7) which will be rendered into the (ψ, θ) coördinates by the foregoing analysis presently. In order to avoid the one degree of arbitrariness existing in this problem we assume the solution of ν to be

(7.10) $$g_{44} \equiv e^{2\nu} = 1 - 2M\psi$$

where M is the mass of the body and ψ according to (6.9) is the Newtonian potential per unit mass. This general assumption includes apparently Schwarzschild's solution as a special case. Equation (7.9) then becomes

(7.11) $$\partial n/\partial \psi - \partial g/\partial \psi - \partial \nu/\partial \psi = 0,$$

giving

(7.12) $$e^\gamma = e^{\nu+\lambda-\mu}\Theta(\theta)$$

where $\Theta(\theta)$ is an arbitrary function. At infinity where $\psi = 0$, $\lambda = \mu = \gamma$

$=\nu=0$ for all values of θ. Hence we have $\Theta(\theta)\equiv 1$ and (7.12) can be rewritten in the form,

$$\lambda+\nu=\gamma+\mu. \tag{7.13}$$

This condition is evidently satisfied by Schwarzschild's solution and even by the plane (5.12).

By (7.10) and (7.11) and since $\Phi=e^{n+\nu}$, (7.8) becomes

$$2\frac{\partial}{\partial\psi}\left[(1-2M\psi)\frac{\partial n}{\partial\psi}\right]+\frac{\partial^2}{\partial\theta^2}e^{2n}=0. \tag{7.14}$$

Consider (2.6). By (2.2), the exact differential

$$dm=m_u du+m_v dv \tag{7.15}$$

can be integrated into the form,

$$2m=\log(\chi_u^2+\chi_v^2)+\chi+2\int Pdu+\theta dv, \quad \text{where} \tag{7.16}$$

$$P=\chi_v C+\chi_u D,\quad Q=\chi_u C-\chi_v D, \tag{7.17}$$

$$\begin{aligned} C&=(\chi_u^2+\chi_v^2)^{-1}(n_u n_v+\nu_u\nu_v),\\ D&=\tfrac{1}{2}(\chi_u^2+\chi_v^2)^{-1}(n_u^2+\nu_u^2-n_v^2-\nu_v^2). \end{aligned} \tag{7.18}$$

Expression (7.16) contains only first partial derivatives and is simpler than (2.6). By (7.5) and the inverse relations, (7.16) becomes

$$2m=\log(\alpha^2+\beta^2)+\log(\chi_\psi^2+e^{2g}\chi_\theta^2)+\chi+2\int P'd\psi+Q'e^{-g}d\theta; \tag{7.19}$$

$$\begin{aligned} P'&=e^g\chi_\theta C'+\chi_\psi D',\quad Q'=\chi_\psi C'-e^g\chi_\theta D\ ;\\ C'&=(\chi_\psi^2+e^{2g}\chi_\theta^2)^{-1}e^g(n_\psi n_\theta+\nu_\psi\nu_\theta),\\ D'&=\tfrac{1}{2}(\chi_\psi^2+e^{2g}\chi_\theta^2)^{-1}(n_\psi^2+\nu_\psi^2-e^{2g}\,[n_\theta^2+\nu_\theta^2]). \end{aligned} \tag{7.20}$$

Between (7.4) and (7.19) we can eliminate the auxiliary functions m and $\alpha^2+\beta^2$. A recapitulation of results gives

$$2\lambda=\log(\chi_\psi^2+e^{2g}\chi_\theta^2)+2H+\chi+2\int P'd\psi+Q'e^{-g}d\theta, \tag{7.21}$$

$$g_{44}\equiv e^{2\nu}=1-2M\psi, \tag{7.10}$$

$$\nabla^2\psi=0, \tag{6.9}$$

$$\lambda+\nu=\gamma+\mu, \tag{7.13}$$

$$2\frac{\partial}{\partial\psi}\left[(1-2M\psi)\frac{\partial n}{\partial\psi}\right]+\frac{\partial^2}{\partial\theta^2}e^{2n}=0, \tag{7.14}$$

$$e^n=\rho e^\gamma,\quad \chi=n+\nu,\quad e^g=\rho e^{\lambda-\mu}=\rho e^{\gamma-\nu},$$

$$(6.12)\quad ds^2 = e^{-2H}\,[e^{2\lambda}d\psi^2 + \rho^{-2}e^{2\mu}d\theta^2] - \rho^2 e^{2\gamma}d\phi^2 + e^{2\nu}dt^2.$$

$$(6.13)\quad e^{-2H} = h^2(\psi\ ^2 + \psi\ ^2)^{-1}.$$

From the above list we see immediately that the solution of the whole problem depends upon the solution of the well known Laplace's equation, (6.9), the non-linear equation, (7.14) and a quadrature, (7.21).

IV. Fields of Spheroidal Homoeoids.

8. *Oblate spheroidal homoeoid.* Let the equation of the homoeoid be

$$(8.1)\qquad \rho^2/a^2 + z^2/c^2 = 1, \qquad (\rho^2 = x^2 + y^2,\quad a^2 > c^2).$$

Use spheroidal coördinates, ξ, η, defined by

$$(8.2)\qquad \rho + iz = \kappa\cos(\xi + i\eta) \qquad (\kappa^2 = a^2 - c^2).$$

Then $\eta =$ const. represents a family of oblate spheroids confocal with (8.1), which is $\kappa\cosh\eta = a$ in the family and $\xi =$ const. a family of hyperboloids of one sheet confocal with and orthogonal to the spheroids.

The Newtonian potential for an oblate spheroidal homoeoid with unit mass is

$$(8.3)\qquad \psi = \kappa^{-1}\cot^{-1}(\sinh\eta).$$

The function, θ, defined by (6.7) may be taken as

$$(8.4)\qquad \theta = \sin\xi.$$

From (8.2), (8.3) and (8.4),

$$(8.5)\qquad \rho^2 = \kappa^2(1 - \theta^2)\csc^2\kappa\psi.$$

In the present case (6.12) is

$$(8.6)\quad ds^2 = -\,e^{-2H}\,[e^{2\lambda}d\psi^2 + \rho^{-2}e^{2\mu}d\theta^2] - \rho^2 e^{2\gamma}d\phi^2 + e^{2\nu}dt^2,$$

where
$$e^{-2H} = \kappa^4\cosh^2\eta(\sinh^2\eta + \sin^2\xi),$$

and λ, μ, γ are to be determined, ν being given by (7.10).

The equation (7.14) that γ must satisfy becomes

$$(8.7)\qquad \frac{d}{d\psi}\left[(1 - 2M\psi)\left(\frac{d\gamma}{d\psi} - \kappa\cot\kappa\psi\right)\right] - \kappa^2 e^{2\gamma}\csc^2\kappa\psi = 0,$$

in which we assume that γ is a function of ψ alone. Equation (8.7) is solvable by the following changes of variables,

$$(8.8)\quad R = e^{\gamma}(1 - 2M\psi)^{1/2}\csc\kappa\psi, \quad du = -\csc\kappa\psi\cdot e^{\gamma}(1 - 2M\psi)^{-1/2}\kappa\,d\psi,$$

where the negative sign in du is chosen to make u tend toward positive infinity as ψ approaches zero. (8. 7) then becomes

$$d^2R/du^2 - R = 0, \quad \text{giving} \tag{8.9}$$

$$R = c_1 \sinh u. \tag{8.10}$$

By (8. 8) and (8. 10), we have

$$c_1 \sinh u = e^{\gamma}(1 - 2M\psi)^{1/2} \csc \kappa\psi \tag{8.11}$$

from which, and the choice of du in (8. 8), we see that c_1 must be a positive constant. Eliminating e^{γ} between du in (8. 8) and (8. 11) we get

$$-c_1\kappa(1 - 2M\psi)^{-1}\, d\psi = \operatorname{csch} u\, du, \quad \text{which gives} \tag{8.12}$$

$$(c_1\kappa/2M) \log (1 - 2M\psi) = \log (\coth u - \operatorname{csch} u) + c_2. \tag{8.13}$$

Eliminating u between (8. 11) and (8. 13) we find

$$\begin{aligned} \gamma = & -\tfrac{1}{2}(c_1\kappa/M + 1) \log (1 - 2M\psi) \\ & + \log \{[e^{2\gamma}(1 - 2M\psi) + c_1^2 \sin^2 \kappa\psi]^{1/2} - c_1 \sin \kappa\psi\} + c_2, \end{aligned} \tag{8.14}$$

the general solution of (8. 7) involving the two constants, c_1 and c_2. To determine c_2 let ψ approach zero. Then γ tends toward zero and $c_2 = 0$. The constant c_1 can be identified with M/k, which is also a constant of integration in Newton's theory. (8. 14) now becomes

$$e^{\gamma} = k^{-1}\psi^{-1} \sin k\psi. \tag{8.15}$$

Obviously e^{γ} approaches unity as ψ tends toward zero.

Last, we must obtain λ. Knowing λ we can get μ by (7. 13) and (8. 15). In order to avoid cumbersome differentiations in integrating (7. 21) directly we use the transformation of ψ in (8. 8), and furthermore put $dv = d\xi$. By (7. 13), (8. 3), (8. 5), (8. 8) and (8. 15), (8. 6) can be written in the form,

$$ds^2 = -\kappa^{-2}e^{-2H}e^{2\mu} \sin^2 \kappa\psi\, [du^2 + dv^2] - \psi^{-2} \cos^2 v\, d\phi^2 + e^{2\nu}dt^2, \tag{8.16}$$

which has the same form as (1. 4), provided

$$e^{2m} = \kappa^2 e^{2\mu}(\cot^2 \kappa\psi + \sin^2 v). \tag{8.17}$$

By (8. 8) and (8. 15), we see that ψ can be expressed as an explicit function of u, and (2. 6), that must be satisfied by μ, can be computed with the aid of R in (8. 10). The quadrature in terms of the u, v variables is quite simple. Coupled with the condition that at infinite distances from the body μ must vanish, the function, μ, is found to be

$$e^{2\mu} = \kappa^{-2}\psi^{-2}(\sinh^2 \eta + \sin^2 \xi)^{-1}. \tag{8.18}$$

From (7.13) and (8.18), λ is given by

$$(8.19) \qquad e^{2\lambda} = \kappa^{-4}\psi^{-4}\operatorname{sech}^2\eta(1-2M\psi)^{-1}(\sinh^2\eta+\sin^2\xi)^{-1}.$$

Again, ds^2 in (8.6) becomes

$$(8.20) \qquad \begin{aligned} ds^2 = &-\kappa^{-2}\psi^{-2}[\psi^{-2}(1-2M\psi)^{-1}\operatorname{sech}^2\eta d\eta^2 \\ &+\kappa^2 d\xi^2+\kappa^2\cos^2\xi d\phi^2]+(1-2M\psi)dt^2. \end{aligned}$$

Solving for $\sinh\eta$ and $\sin\xi$ from (8.2) we get

$$(8.21) \qquad \begin{aligned} 2\kappa^2\sinh^2\eta &= r^2-\kappa^2+[r^4-2\kappa^2(\rho^2-z^2)+\kappa^4]^{1/2}, \\ 2\kappa^2\sin^2\xi &= -(r^2-\kappa^2)+[r^4-2\kappa^2(\rho^2-z^2)+\kappa^4]^{1/2}. \end{aligned} \qquad (r^2=\rho^2+z^2)$$

When κ is small compared with r, these expressions can be expanded in the following forms:

$$(8.22) \qquad \begin{aligned} \kappa^2\sinh^2\eta &= r^2\left[1-\frac{\rho^2}{r^4}\kappa^2+\frac{(1-\omega^2)}{4r^4}\kappa^4+\cdots\right], \\ \sin^2\xi &= \frac{z^2}{r^2}+\frac{1-\omega^2}{4r^2}\kappa^2+\frac{\omega(1-\omega^2)}{4r^4}\kappa^4+\cdots, \\ \omega &= (\rho^2-z^2)/(\rho^2+z^2). \end{aligned}$$

It is interesting to observe from (8.22) that when κ approaches zero, namely, when the spheroidal homoeoid tends toward a spherical shell as a limit, the line element (8.20) becomes Schwarzschild's solution. Furthermore, (8.20) is also the solution of an infinitely thin material disc with mass M and radius κ.

9. *Prolate spheroidal homoeoid.* The treatment of the prolate spheroidal homoeoid is analogous to the preceding problem. Here the equation of the surfaces of the body is given by

$$(9.1) \qquad \rho^2/a^2+z^2/c^2=1 \quad \text{with} \quad c^2>a^2.$$

The spheroidal coördinates ξ, η, used are defined by

$$(9.2) \qquad z+i\rho=\kappa\cos(\xi+i\eta), \qquad (\kappa^2=c^2-a^2).$$

The Newtonian potential for (9.1) with unit mass is

$$(9.3) \qquad \psi=\frac{1}{2\kappa}\log\frac{\cosh\eta+1}{\cosh\eta-1}.$$

The function, θ, defined by (6.7) becomes

$$(9.4) \qquad \theta=-\cos\xi,$$

and ds^2 given in (6.12) is then

$$(9.5)\qquad \begin{aligned} ds^2 &= -\,e^{-2H}\left[e^{2\lambda}d\psi^2 + \rho^{-2}e^{2\mu}d\theta^2\right] - \rho^2 e^{2\gamma}d\phi^2 + e^{2\nu}dt^2,\\ e^{-2H} &= \kappa^4 \sinh^2\eta\,(\sinh^2\eta + \sin^2\xi),\quad e^{2\nu} = 1 - 2M\psi, \end{aligned}$$

where γ is assumed to be a function of ψ alone while both λ and μ are functions of ψ and θ. Between $\lambda, \mu, \gamma, \nu$ we have the relation (7.13), namely $\lambda + \nu = \gamma + \mu$. The equation (7.14) for γ in the present case is similar to (8.7), so the remaining analysis will be similar. The result is

$$(9.6)\qquad ds^2 = -\,\kappa^2(\sinh^2\eta + \sin^2\xi)\left[e^{2\lambda}d\eta^2 + e^{2\mu}d\xi^2\right] - \rho^2 e^{2\gamma}d\phi^2 + e^{2\nu}dt^2,$$

where
$$\begin{aligned} e^{2\lambda} &= \kappa^{-4}\psi^{-4}\operatorname{csch}^2\eta\,(1-2M\psi)^{-1}(\sinh^2\eta + \sin^2\xi)^{-1},\\ e^{2\mu} &= \kappa^{-2}\psi^{-2}\left[\sinh^2\eta + \sin^2\xi\right]^{-1},\\ e^{2\gamma} &= \kappa^{-2}\psi^{-2}\operatorname{csch}^2\eta. \end{aligned}$$

We notice that (9.6) is also the solution for a rod of length κ and mass M lying on the z-axis. Similarly when the prolate spheroidal homoeoid approaches a spherical shell as a limit, (9.6) degenerates into Schwarzschild's solution.

10. *Motion of a particle in the field of an oblate spheroidal homoeoid.* The fundamental quadratic differential form (8.20) for an oblate spheroidal homoeoid can also be written in the form

$$(10.1)\qquad ds^2 = -\,\psi^{-4}(1-2M\psi)^{-1}d\psi^2 - \psi^{-2}d\xi^2 - \psi^{-2}\cos^2\xi\,d\phi^2 + (1-2M\psi)\,dt^2.$$

If for convenience we put

$$(10.2)\qquad \psi = 1/r,\quad \xi = \theta - \pi/2,$$

where it must be remembered that r and θ are not the r and θ used in previous sections, then (10.1) becomes

$$(10.3)\qquad ds^2 = -(1-2M/r)^{-1}dr^2 - r^2d\theta^2 - r^2\sin^2\theta\,d\phi^2 + (1-2M/r)\,dt^2,$$

which has the same analytical form as Schwarzschild's solution. The results worked out in the latter case are immediately applicable to the present problem, provided we interpret the symbols in (10.3) appropriately.

The four differential equations,

$$(10.4)\qquad \frac{d^2x_\alpha}{ds^2} + \{\mu\nu,\alpha\}\frac{dx_\mu}{ds}\frac{dx_\nu}{ds} = 0,$$

defining the motion of an infinitesimal particle in the four dimensional continuum characterized by (10.3) are *

* A. S. Eddington, *loc. cit.*, pp. 85.

$$(10.5)\qquad \frac{d^2r}{ds^2}+\lambda'\left(\frac{dr}{ds}\right)^2-re^{-2\lambda}\left(\frac{d\theta}{ds}\right)^2-r\sin^2\theta e^{-2\lambda}\left(\frac{d\phi}{ds}\right)^2 + e^{2\nu-2\lambda}\nu'\left(\frac{dt}{ds}\right)^2=0,$$

$$(10.6)\qquad \frac{d^2\theta}{ds^2}+\frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds}-\sin\theta\cos\theta\left(\frac{d\phi}{ds}\right)^2=0,$$

$$(10.7)\qquad \frac{d^2\phi}{ds^2}+\frac{2}{r}\frac{dr}{ds}\frac{d\phi}{ds}+2\cot\theta\frac{d\theta}{ds}\frac{d\phi}{ds}=0,$$

$$(10.8)\qquad \frac{d^2t}{ds^2}+2\nu'\frac{dr}{ds}\frac{dt}{ds}=0,$$

where $e^{2\nu}=1-2M/r,\quad \lambda+\nu=0,\quad \nu'=d\nu/dr.$

Instead of using (10.5) we can take (10.3), which can be written as

$$(10.9)\qquad e^{-2\nu}\left(\frac{dr}{ds}\right)^2+r^2\left(\frac{d\theta}{ds}\right)^2+r^2\sin^2\theta\left(\frac{d\phi}{ds}\right)^2-e^{2\nu}\left(\frac{dt}{ds}\right)^2=-1.$$

Equations (10.7) and (10.8) are immediately integrable, giving respectively

$$(10.10)\qquad r^2\sin^2\theta(d\phi/ds)=c_2,$$

$$(10.11)\qquad dt/ds=c_1e^{-2\nu}.$$

Let the constants of integration c_1 and c_2 be positive.

Eliminating $d\phi/ds$ between (10.6) and (10.10), we find

$$(10.12)\qquad \frac{d^2\theta}{ds^2}+\frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds}-\frac{c_2{}^2}{r^4}\cos\theta\csc^3\theta=0,\quad \text{giving}$$

$$(10.13)\qquad r^4(d\theta/ds)^2+c_2{}^2\csc^2\theta=c_3{}^2\qquad (\text{Take } c_3>0).$$

Eliminating ds from (10.10) and (10.13), we get

$$(10.14)\qquad d\phi=-c_2\,[(c_3{}^2-c_2{}^2)-c_2{}^2\cot^2\theta]^{-\frac{1}{2}}\csc^2\theta d\theta,$$

in which we choose the negative sign to make θ decrease when ϕ increases. Let

$$(10.15)\qquad p=(c_3{}^2-c_2{}^2)^{\frac{1}{2}}/c_2.$$

By (10.13) since r, θ, s are all real we see that $c_3{}^2\geqq c_2{}^2$ and consequently p is real. Integrating (10.14), we obtain

$$(10.16)\qquad \cot\theta=p\sin(\phi-\Omega),$$

where Ω is the node, and θ is taken to be $\pi/2$ when $\phi=\Omega$. The geo-

metrical meaning of $\theta=\pi/2$ is that $z=0$ where the particle crosses the equatorial plane of the oblate homoeoid [cf. (10.2) and (8.2)].

By using (10.9), (10.10), (10.11), (10.13) and (10.16), we obtain the following relation between ψ and ϕ,

(10.17) $$c_2[2Mf(\psi)]^{-\frac{1}{2}}d\psi = -c_3[1+p^2\sin^2(\phi-\Omega)]^{-1}d\phi,$$ where

(10.18) $$f(\psi)\equiv\psi^3-\frac{1}{2M}\psi^2+\frac{1}{c_3^2}\psi+\frac{c_1^2-1}{2Mc_3^2};$$

the negative sign in (10.17) will be explained presently.

The right hand side of (10.17) is immediately integrable in terms of circular functions. The rigorous integration of the left hand side in terms of elliptic functions has been discussed by Forsyth.* Let α, β, γ $(\alpha>\beta>\gamma)$ be the three roots of $f(\psi)$. Then ψ can lie only with the interval $\beta\geqq\psi\geqq\gamma$. When $\psi=\beta$, we have the analogous "perihelion" and when $\psi=\gamma$, the "aphelion." Let $\phi=\phi_0$, when $\psi=\beta$. Integrating, we have

(10.19) $$c_2\int_\beta^\psi[2Mf(\psi)]^{-\frac{1}{2}}d\psi = -c_3\int_{\phi_0}^\phi[1+p^2\sin^2(\phi-\Omega)]^{-1}d\phi.$$

Here we see that since ψ decreases after $\psi=\beta$, but that ϕ continues to increase after $\phi=\phi_0$, the negative sign in (10.17) must be taken.

From (10.19) we obtain

(10.20) $$\psi=\gamma+(\beta-\gamma)(1-cn2\mu)/(1+dn2\mu),$$

where μ is defined by the equation,

(12.21) $$2\mu=2K-(1/P)\{\tan^{-1}[\sigma\tan(\phi-\Omega)]-\tan^{-1}[\sigma\tan(\phi_0-\Omega)]\},$$

in which $$\sigma=c_3/c_2,\quad P=[2M(\alpha-\gamma)]^{-\frac{1}{2}},$$

and K is the complete elliptic integral of the first kind with modulus, k, given by

(10.22) $$k^2=(\beta-\gamma)/(\alpha-\gamma).$$

From (8.2), (8.3), (10.2), (10.16), (10.20) and (10.21), we obtain the equations of the orbit of the particle in the following forms,

(10.23) $$\rho=\kappa[1+p^2\sin^2(\phi-\Omega)]^{-\frac{1}{2}}\csc\kappa\psi,$$ $$z=\kappa p\sin(\phi-\Omega)[1+p^2\sin^2(\phi-\Omega)]^{-\frac{1}{2}}\cot\kappa\psi.$$

The equation $\theta=$ const. [cf. (10.2) and (8.2)] represents the family of hyperboloids of one sheet orthogonal to the family of spheroids $\psi=$ const.

---

* A. R. Forsyth, *Proceedings of the Royal Society* (A), Vol. 97 (1920), pp. 145.

Then (10.16) shows that the maximum and minimum latitudes of the particle in its orbit are invariable for given initial conditions.

The function, ψ, in (10.20) is a Jacobian elliptic function of ϕ. Hence the analogous "line of apsides" of the orbit precesses about the z axis. The amount of this precession for the particle to describe the orbit once can be calculated in the following manner.* In (10.20) we have so chosen ψ, ϕ that at perihelion $\psi = \beta$, $\phi = \phi_0$. Then at aphelion $\psi = \gamma$, let $\phi = \phi_1$. From (10.20) and (10.21),

$$(10.24) \quad \tan^{-1}[\sigma \tan(\phi_1 - \Omega)] - \tan^{-1}[\sigma \tan(\phi_0 - \Omega)] = 2PK.$$

At the next perihelion let $\phi = \phi_2$. The relation analogous to (10.24) is

$$(10.25) \quad \tan^{-1}[\sigma \tan(\phi_2 - \Omega)] - \tan^{-1}[\sigma \tan(\phi_1 - \Omega)] = 2PK.$$

Adding (10.24) and (10.25), we get

$$(10.26) \quad \tan^{-1}[\sigma \tan(\phi_2 - \Omega)] - \tan^{-1}[\sigma \tan(\phi_0 - \Omega)] = 4PK.$$

The precession is given by

$$(10.27) \qquad \Delta = \phi_2 - \phi_0 - 2\pi.$$

Solving ϕ_2 from (10.26), we get

$$(10.28) \quad \Delta = \tan^{-1} \frac{1}{\sigma} \left\{ \frac{\tan 4PK + \sigma \tan(\phi_0 - \Omega)}{1 - \sigma \tan(\phi_0 - \Omega) \tan 4PK} \right\} - (\phi_0 - \Omega) - 2\pi.$$

It is interesting to observe from (10.13), (10.15) and (10.23) that if the particle lies initially in the equatorial plane of the homoeoid, i. e. $d\theta/ds = 0$ when $\theta = \pi/2$, then subsequently $\theta = \pi/2$ and the particle will continually lie there. The approximate formula for Δ in this case can be calculated as follows: Regard (r, ϕ) as configurational polar coördinates of the particle. Then (10.3) shows that the motion of the particle in these coördinates is the same as the motion of a corresponding particle in Schwarzschild's solution. Hence the constants, c_1 and c_2, in (10.11) and (10.10) are given by †

$$(10.29) \quad c_2{}^2 = r_0(1 - e^2)M, \quad c_1{}^2 - 1 = -M/r_0, \quad e^2 = (r_0{}^2 - r_1{}^2)/r_0{}^2,$$

where M is the mass of the homoeoid, r_0 the semi-major axis, r_1 the semi-minor axis, and e the eccentricity of the orbit in the configurational coördinate system. The advance of the perihelion is given approximately by

$$(10.30) \qquad \Delta = 2\pi \cdot 3M/r_0(1 - e^2).$$

* A. R. Forsyth, *loc. cit.*, p. 148.

† A. R. Forsyth, *loc. cit.*, p. 145.

From (8.3) and (8.21) with $z=0$, and (10.2), we obtain

$$\frac{1}{r}=\psi=\frac{1}{\kappa}\cot^{-1}\left[\frac{1}{\kappa}(\rho^2-\kappa^2)^{1/2}\right] \tag{10.31}$$

When $\rho^2-\kappa^2>\kappa^2$ which is obviously satisfied by large values of ρ, we can expand (10.31) in ascending powers of κ/ρ in the form,

$$\frac{1}{r}=\frac{1}{\rho}\left[[1+\frac{1}{6}\frac{\kappa^2}{\rho^2}+\frac{3}{40}\frac{\kappa^4}{\rho^4}+\cdots].\right. \tag{10.32}$$

Equation (10.31) shows that ρ is a monotonic function of r, and consequently the value of Δ in (10.30), which is primarily for the orbit in the (r,ϕ) configurational coördinates will hold also in the (ρ,ϕ) system. Knowing the "semi-major" and "semi-minor" axes, ρ_0 and ρ_1 of the particle's orbit in the latter system we can compute the corresponding values of r_0 and r_1 by (10.32). Then (10.30) shows that the oblateness of the central body causes a small increase in the advance of the perihelion of the orbit predicted from Schwarzschild's solution. This increase vanishes when $\kappa=0$, namely, when the oblate spheroidal homoeoid degenerates into a spherical shell.

In conclusion I wish to thank Prof. E. T. Bell heartily, whose interest in this problem and encouragement in the course of work have made this investigation possible.

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIFORNIA.

ON THE NON-EXISTENCE OF CURVES OF ORDER 8 WITH 16 CUSPS.

By Oscar Zariski.

1. In a paper, which will be published in the "Annals of Mathematics", I prove incidentally by means of several examples, that the answer to the well-known question,* as to whether there always exist plane algebraic curves with assigned Plückerian characters, is negative. It is understood, of course, that the assigned characters satisfy the Plücker relations, and are essentially non-negative integers. For instance, I prove that a curve of order 7 cannot possess 11 (or more) cusps.† Thus, with the usual notations for the Plückerian characters of a curve (n-order, m-class, d-number of nodes, k-number of cusps, τ-number of double tangents, i-number of flexes), we have then that the following sets of Plückerian characters do not correspond to any effectively existent curve:

$$(1) \quad n=7,\ d=0,\ k=11;\ m=9,\ \tau=7,\ i=17;$$

$$(2) \quad n=m=7,\ d=\tau=1,\ k=i=11.$$

Similarly, I prove that a curve of order 8 cannot possess 16 (or more) cusps. For instance, there do not exist (self-dual) curves of order 8 with 16 cusps and no nodes.‡

The proof of the non-existence of the above curves is based on the following general theorem, proved in my paper as a consequence of certain results concerning the determination of irregular cyclic multiple planes: *the linear*

---

* See the interesting paper by S. Lefschetz, "On the Existence of Loci with Given Singularities," *Transactions of the American Mathematical Society*, Vol. 14 (1913), pp. 23-41, which constitutes perhaps the first attempt to throw light on questions as yet unsolved in all their generality, relative to the *dimension* and *existence* of continuous systems of cuspidal curves. The examples given below solve some of the questions explicitly raised by Lefschetz in his paper.

† On the other hand, seventics with 10 cusps exist, as is shown by the example of the seventic with 10 cusps and 3 nodes, which is the dual of a sextic curve with 7 cusps and one node.

‡ In contradiction with a statement made by B. Segre in his paper "Esistenza e dimensione di sistemi continui di curve piane algebriche con dati caratteri," *Rendiconti della R. Accademia Nazionale dei Lincei*, (July, 1929), p. 38. The above examples correspond to the lowest possible values of n. In fact, it is easily shown that when $n \leqq 6$, any set of (non-negative) Plückerian characters belong to some effectively existent curve. For instance, the existence of a curve of order 6 with 7 cusps and one node, mentioned in the previous footnote, can be ascertained as follows: The dual of a quartic curve with two cusps is a curve of order 6 possessing 8 cusps and one node. Since the complete continuous system of curves, having the same characters, has its characteristic series non-special, one of the 8 cusps can be considered as virtually non-existent.

system of curves of order $n-3-j$ determined by simple basis points at the cusps of a plane irreducible algebraic curve of order n, is regular (effective dimension = virtual dimension) *for any value of the integer j, such that* $6j < n$. The non-existence of the above mentioned cuspidal curves follows at once from this theorem, if we observe that in each of these examples the corresponding curves, if they existed, would be certainly irreducible and that in each case one of the linear systems, mentioned in the above theorem, would have a virtual dimension equal to -2.

In my paper the quoted theorem is based on considerations, which involve the elements of the theory of algebraic surfaces. Since the question of the Plückerian characters of a curve, so elementary as far as its formulation is concerned, is essentially a problem of plane geometry, it seems desirable to consider one of the examples quoted above and to arrive at the desired conclusion of the non-existence in a more elementary way, by making use only of the elements of plane geometry and of the geometry on an algebraic one-dimensional variety. Accordingly, we propose to give in this paper a direct proof of the non-existence of curves of order 8 with 16 cusps.

2. For the proof, let us suppose that a curve,

$$f(x,y)=0,$$

of order 8 with 16 cusps and no nodes exists. This curve is self-dual, i. e., $m=8$, $\tau=0$, $i=16$. We shall have to consider the curve f from its dual aspect, and in doing this we shall face the question whether the dual singularities of f are exactly of the same type as the point singularities of f, namely ordinary flexes and no double tangents. This is by no means obvious. The general question involved is the following: given a curve f possessing cusps and nodes only, what can be said about the nature of the dual singularities of f? Obviously, it would not be correct to say, that the dual singularities of f are necessarily double tangents and ordinary flexes. Thus a curve without point singularities at all can possess flexes of order $s>1$ (where the flex tangent has $s+2$ coincident intersections with the curve), triple tangents, etc. The question raised must refer, in its correct form, not to the individual curve f, but to the most general curve of the same order and having the same singularities as f, or better, to the generic curve of a complete continuous system of such curves. We can ask namely, whether the dual singularities of the *generic* curve of a complete continuous system of curves, having cusps and nodes only, are exactly double tangents and ordinary flexes. Obvious as the answer may seem, it requires a proof. However, it is not our intention to attempt such a proof in this paper. For our purpose it will be sufficient to show that in one particular case the answer to the above question is in the

affirmative, and not only for the generic curve of the system but for *every* individual curve of the system possessing nodes and cusps only. This is the case in which $\tau = 0$. We propose then to prove the following

LEMMA. *Given an irreducible curve f of order n, possessing d nodes and k cusps. If the number τ of the double tangents, evaluated according to the Plücker relations, has the value* 0, *then the only dual singularities of f are exactly ordinary flexes.*

Denote, as usual, by i the number of flexes, as evaluated according to the Plücker formulas, and by p the genus of f. We shall have

$$p = (n-1)(n-2)/2 - d - k = (m-1)(m-2)/2 - i,$$

where m is the class of f. Denote by $\bar{f}$ the transform of f by duality. Let $q_1, q_2, \cdots, q_l$ be the singular tangents of f, and let $Q_1, Q_2, \cdots, Q_l$ be the corresponding singular points of $\bar{f}$. The line q_j will touch f at certain points, having at each point a contact of a certain order with f. In addition q_j may pass through some of the nodes and cusps of f. To each contact of order s will correspond by duality a branch of $\bar{f}$ of order s and of origin Q_j, in the neighborhood of which there will be only simple points infinitely near to Q_j. To distinct points of contact there will correspond branches with distinct tangents. If q_j passes through a node, we shall have an additional branch of origin Q_j, if and only if q_j coincides with one of the principal tangents of the double point. Similarly, if q_j passes through a cusp, we shall have an additional linear branch of origin Q_j, if and only if q_j coincides with the cusp tangent.

Let $s_{j1}, s_{j2}, \cdots, s_{jr_j}$ be the orders of contact of q_j at the r_j points at which it touches f, including possibly the nodes and cusps, through which q_j happens to pass and at which it coincides with a principal tangent.*

Let

$$s_j = s_{j1} + s_{j2} + \cdots + s_{jr_j}.$$

Then Q_j is an s_j-fold point of $\bar{f}$, and since in its neighborhood there are only simple points it follows that Q_j absorbs $s_j(s_j-1)/2$ ordinary double points of $\bar{f}$. Hence

$$p = (m-1)(m-2)/2 - \sum_{j=1}^{l} s_j(s_j-1)/2,$$

and consequently

$$\sum_{j=1}^{l} s_j(s_j-1)/2 = i. \tag{1}$$

* At any of these cusps we must put $s = 1$. At a node the order of contact of q_j with one of the branches through the double point is to be considered.

We also have,

$$n = m(m-1) - 3i.$$

On the other hand it is obvious that the multiple point Q_j of $\bar{f}$ diminishes the class of $\bar{f}$ by

$$s_j(s_j-1) + (s_{j1}-1) + (s_{j2}-1) + \cdots + (s_{jr_j}-1) = s_j^2 - r_j.$$

Hence

$$n = m(m-1) - \sum_{j=1}^{l} (s_j^2 - r_j),$$

and consequently

$$\sum_{j=1}^{l} (s_j^2 - r_j) = 3i. \tag{2}$$

From (1) and (2) we deduce:

$$3 \sum_{j=1}^{l} s_j(s_j-1)/2 = \sum_{j=1}^{l} (s_j^2 - r_j),$$

or

$$\sum_{j=1}^{l} [s_j(s_j-3)/2 + r_j] = 0. \tag{3}$$

Since for any j, $s_j \geqq 2$ and r_j is positive, it follows that each term of the written sum is $\geqq 0$, where the sign $=$ holds, if and only if $s_j = 2$ and $r_j = 1$. Hence the relation (3) implies that $s_j = 2$, $r_j = 1$ for $j = 1, 2, \cdots, l$, i. e., that the only singularities of $\bar{f}$ are ordinary cusps, which proves our Lemma.

COROLLARY. *A curve f of order* 8 *with* 16 *cusps and no nodes, if it exists, possesses exactly* 16 *ordinary flexes, i. e., the transform $\bar{f}$ of f will also possess* 16 *ordinary cusps.*

3. Let $f(x, y) = 0$ be a curve (necessarily irreducible) of order 8 with 16 cusps and no nodes. Its genus is 5. Denote by Γ_8 a set of 8 points cut out on f by a line of the plane, by $\bar{\Gamma}_8$ a set of 8 points, outside the cusps, cut out on f by a first polar of f, and by $|\Gamma_8|$ and $|\bar{\Gamma}_8|$ the corresponding complete series on f. Let moreover H_8 be a canonical set on f, i. e., a set of the canonical series g_8^4. Furthermore denote by K_{16} the set of 16 cusps of f. We then have the following relations:

$$2\,\Gamma_8 + H_8 \equiv \bar{\Gamma}_8 + K_{16}; \tag{4}$$

$$5\,\Gamma_8 \equiv H_8 + 2\,K_{16}. \tag{4'}$$

The relation (4) expresses the known fact that the Jacobian set $\Gamma_{24} \equiv \bar{\Gamma}_8 + K_{16}$ of the g_8^1 cut out on f by a pencil of lines is equivalent to the sum of a canonical set and of the two-fold of a set of the g_8^1. The relation (4′) expresses the fact that the canonical series is cut out on f by the adjoint curves of order 5. Eliminating the set K_{16} between (4) and (4′) we obtain

(5) $$\Gamma_8 + 2\,\bar{\Gamma}_8 \equiv 3\,H_8.$$

Since the curve f is self-dual, and since by the Lemma proved above, the curve $\bar{f}$ dual to f is exactly of the same type as f, the relation (5) can be dualized, by interchanging Γ_8 and $\bar{\Gamma}_8$ and by leaving H_8 unaltered. Hence

(5′) $$\bar{\Gamma}_8 + 2\,\Gamma_8 \equiv 3\,H_8.$$

From (5) and (5′) we deduce $\Gamma_8 \equiv \bar{\Gamma}_8$, i. e., the two series $|\,\Gamma_8\,|$ and $|\,\bar{\Gamma}_8\,|$ coincide.

There are now two possible cases to consider: (1) the series $|\,\Gamma_8\,|$ is non-special, and hence is a $g_8{}^3$; (2) the series $|\,\Gamma_8\,|$ is special, and then it necessarily coincides with the canonical series $g_8{}^4 = |\,H_8\,|$. We investigate the two cases separately.

4. Let us first suppose that $|\,\Gamma_8\,| = |\,\bar{\Gamma}_8\,|$ is a non-special series $g_8{}^3$. Let $g_8{}^2$ and $\bar{g}_8{}^2$ denote the (incomplete) series cut out on f by the lines of the plane and by the first polars of f respectively. Since the two series, of dimension 2, are both contained in the series $|\,\Gamma_8\,|$ of dimension 3, they have a $g_8{}^1$ in common. This $g_8{}^1$ is cut out on f by a pencil of lines, say of center A, and on the other hand the same $g_8{}^1$ is cut out on f by a pencil of first polars; the pole B of the variable polar of the pencil will describe a line b. We have then the following situation: there exsts a point A and line b in the plane of f, such that for any point B of the line b the points of contact with f of the 8 tangents drawn from B are the intersections with f of a line a on A. It is easily seen that this situation is impossible. For, let us consider a point B_1 at which the line b meets the curve f. If A happens to be on b and on f, we may suppose that B_1 is distinct from A, since f possesses only ordinary flexes. Let a_1 be the line on A which corresponds to B_1. Since, as the variable point B of b approaches B_1, two or more of the points of contact with f of the tangents drawn from B approach B_1, it follows that the line a_1 necessarily coincides with the line AB_1. It follows that the line a_1 must absorb all the 8 tangents, which can be drawn through B_1, which is impossible, since f possesses only ordinary flexes.

5. We now consider the case, in which the series $|\,\Gamma_8\,| = |\,\bar{\Gamma}_8\,|$ coincides with the canonical series $g_8{}^4$. In this case there exist adjoint curves of order 5, which cut out on f sets of the $g_8{}^2$ cut out by the lines of the plane. These adjoint curves necessarily degenerate into a line and into a fixed quartic curve. Hence the 16 cusps of f lie on a quartic curve, which obviously does not meet f outside the cusps. It can also be shown that there exists a sextic curve, which passes through the cusps of f and touches at each cusp the cusp tangent (and which therefore does not meet f outside the cusps). In fact, the relations (4) and (4′) yield by subtraction:

$$K_{16} \equiv 3\,\Gamma_8 + \bar{\Gamma}_8 - 2\,H_8,$$

or, in view of (5′),

$$K_{16} \equiv \Gamma_8 + H_8.$$

This relation shows that the set K_{16} of the cusps of f belongs to the series $|\,\Gamma_8 + H_8\,|$, which is cut out on f by the adjoint curves of order 6. It follows that there exists an adjoint sextic curve, whose 16 intersections with f (generally outside the cusps) fall at the cusps of f. Such a sextic must touch at each cusp the cusp tangent.

Let $\psi_6(x, y) = 0$ and $\phi_4(x, y) = 0$ be the equations of this sextic curve and of the above quartic curve respectively. Let us consider the pencil

$$[\psi_6(x, y)]^2 - t[\phi_4(x, y)]^3 = 0.$$

The curves of this pencil do not have variable intersections with the curve f, since all the intersections fall at the cusps of f. Hence, for a proper value of t, which we may suppose to be $t = 1$, the corresponding curve of the pencil will contain the curve f as a component. We have then

$$[\psi_6(x, y)]^2 - [\phi_4(x, y)]^3 = A_4(x, y) \cdot f(x, y), \tag{6}$$

where $A_4(x, y)$ is a polynominal of order 4 in x and y.

To prove that the relation (6) cannot hold we first observe that the curves $\psi_6 = 0$ and $\phi_4 = 0$, which we shall denote in the sequel by C_6 and C_4, satisfy the following conditions: (1) the cusps of the curve f are *simple* points of the two curves, and at each cusp the two curves have distinct tangents; (2) if the two curves are reducible, they do not have common components; (3) each curve possesses only a finite number of multiple points. In fact, (1) holds, because the two curves C_6 and C_4 have at each cusp of f exactly 3 and 2 coincident intersections with f respectively, otherwise the total number of intersections of one of these curves with f would be greater than the product of its order and the order of f. The condition (2) holds, because a common component of the curves would have to meet f at the cusps only, and the two curves would have at some cusp a common tangent. Finally neither one of the curves can possess a curve of multiple points, because such a curve would have to meet f at the cusps only, which contradicts the condition (1).

It follows that C_6 and C_4 have exactly 8, distinct or coincident, intersections outside of the cusps of f. Each of these intersections is at least a double point of the curve $A_4 = 0$. Thus, if O is a common simple point of the curves C_6 and C_4, then in general O will be a cusp of the curve $A_4 = 0$. If, however, the point O absorbs two intersections of the curves C_6 and C_4, then O is either a tacnode of the second kind or a triple point for the curve $A_4 = 0$, according as C_6 possesses at O a simple point or a double point. In

both cases O will absorb 3 double points of the curve $A_4 = 0$. At any rate the meet points of C_6 and C_4 constitute a set of multiple points of the curve $A_4 = 0$, which will absorb at least 8 double points of the curve. It follows that the curve $A_4 = 0$ possesses *infinite* multiple points, since the maximum number of double points (or of equivalent singularities) which a curve of order n can possess, without possessing infinite multiple points, is $n\ (n-1)/2$. Hence the curve $A_4 = 0$ contains a line of multiple points or is a conic counted twice. We consider in the next two sections these two cases. The reader should bear in mind that the quartic $A_4 = 0$ cannot pass through a cusp of f, and that any point common to the curve $A_4 = 0$ and C_4 (or C_6) is also on C_6 (or C_4), but is not on the curve f.

6. Let the curve $A_4 = 0$ contain a line of multiple points, say the line $x = 0$. Let

$$\psi_6(x, y) = a_6(y) + a_5(y)x + a_4(y)x^2 + a_3(y)x^3 + \cdots,$$
$$\psi_4(x, y) = b_4(y) + b_3(y)x + b_2(y)x^2 + \cdots,$$

where all the missing terms contain higher powers of x. The coefficients $a_i(y), b_j(y)$ are polynominals in y of degrees indicated by the indices. Since, by hypothesis, x^2 is a factor of the polynominal $[\psi_6(x, y)]^2 - [\phi_4(x, y)]^3$, we must have

$$(7) \qquad [a_6(y)]^2 = [b_4(y)]^3;$$
$$(7') \qquad 2a_6(y)a_5(y) = 3[b_4(y)]^2 b_3(y).$$

From (7) we deduce that $a_6(y)$ and $b_4(y)$ are the cube and the square respectively of a polynomial of second degree.* Let the roots of this polynominal be assumed to be $y = 0$ and $y = \eta$. Then

$$(8) \qquad a_6(y) = y^3(y-\eta)^3, \quad b_4(y) = y^2(y-\eta)^2;$$
$$(8') \qquad 2a_5(y) = 3y(y-\eta)b_3(y).$$

We find then

$$(9) \quad \{[\psi_6(x, y)]^2 - [\phi_4(x, y)]^3\}/x^2 = y^2(y-\eta)^2 \{-3/4[b_3(y)]^2 + 2y(y-\eta)a_4(y) - 3y^2(y-\eta)^2 b_2(y)\} + \cdots,$$

where all the missing terms involve the variable x. The points $(0, 0)$ and $(0, \eta)$ are on C_6, C_4 and on the curve $A_4 = 0$, which is made up of the line of double points $x = 0$ and of a residual conic. If $\eta \neq 0$, then, by (9), this residual conic passes through the points $(0, 0)$ and $(0, \eta)$, touching there the

* None of the polynomials $a_6(y)$, $b_4(y)$ can vanish identically, because the identical vanishing of one would imply the identical vanishing of the other, and the line $x = 0$ would be a common component of the two curves C_6 and C_4, which, as we observed above, is impossible.

line $x=0$. If $\eta=0$, the conic has at the point $(0,0)$ at least a 4 point contact with the line $x=0$. In both cases we deduce that this conic must degenerate into two lines, one of which is the line $x=0$, the points of which are therefore at least triple points of the curve $A_4=0$. Hence the left-hand member of (9) must be divisible by x, and we must have

$$(10) \qquad 3/4\,[b_3(y)]^2 = 2y(y-\eta)a_4(y) - 3y^2(y-\eta)^2 b_2(y).$$

If $\eta \neq 0$, then it follows that $b_3(y)$ is divisible by $y(y-\eta)$ and that consequently also $a_4(y)$ is divisible by $y(y-\eta)$. Putting

$$b_3(y) = y(y-\eta)\,c_1(y), \quad a_4(y) = y(y-\eta)\,c_2(y),$$

we have, by (8) and (8′),

$$\begin{aligned}\psi_6(x,y) &= y^3(y-\eta)^3 + \tfrac{3}{2}y^2(y-\eta)^2 c_1(y)x \\ &\qquad + y(y-\eta)c_2(y)x^2 + a_3(y)x^3 + \cdots, \\ \phi_4(x,y) &= y^2(y-\eta)^2 + y(y-\eta)c_1(y)x + b_2(y)x^2 + \cdots.\end{aligned}$$

We see that the curve C_6 possesses at the origin a triple point (at least) and that C_4 possesses at the origin a double point (at least). The origin is therefore at least a 6-fold point of the curve $[\psi_6(x,y)]^2 - [\phi_4(x,y)]^3 = 0$, which is impossible since this would imply that the curve f has at the origin a double point (at least), whereas we know that it does not pass through the origin at all.

Let now $\eta = 0$. Then, by (10), $b_3(y)$ is divisible by y,

$$(11) \qquad b_3(y) = yc_2(y),$$

and (10) becomes

$$(10') \qquad 3/4\,[c_2(y)]^2 = 2a_4(y) - 3y^2 b_2(y).$$

We find

$$\begin{aligned}(12) \quad \{[\psi_6(x,y)]^2 - [\phi_4(x,y)]^3\}/x^3 &= y^3\,\{3c_2(y)a_4(y) \\ &+ 2y^3a_3(y) - [c_2(y)]^3 - 6y^2c_2(y)b_2(y) - 3y^5b_1(y)\} + \cdots,\end{aligned}$$

where all the missing terms involve the variable x. The curve $A_4=0$ is made up of the line of triple points $x=0$, and of a residual line. The presence of the factor y^3 in the right-hand member of the relation (12) shows that this residual line must coincide with the line $x=0$, so that the quartic $A_4=0$ is merely the line $x=0$ counted 4 times. Expressing the fact that the term independent of x in the right-hand member of (12) vanishes identically, we find that $c_2(y)\{3a_4(y) - [c_2(y)]^2\}$ must be divisible by y^2. Since, by (10′), also $3/4\,[c_2(y)]^2 - 2a_4(y)$ is divisible by y^2, we deduce that $a_4(y)$ is divisible by y^2. Taking in account (8), (8′) and (11), we then deduce, as in the previous case ($\eta \neq 0$), that the curves C_6 and C_4 have at the origin at least a triple point and a double point respectively, which is impossible.

7. There remains to consider the case in which the quartic curve $A_4 = 0$ is a conic C_2 counted twice. Let O be a point at which C_2 meets the curve C_4 and which is therefore also on the curve C_6, *but not on the curve f*. Since O is a double point of the curve $[\psi_6(x,y)]^2 - [\phi_4(x,y)]^3 = 0$, it is necessarily a simple point of the curve C_6. It is obvious that C_2 and C_6 have the same tangent at O. We suppose for simplicity that the point O is at the origin, and that the common tangent of C_6 and C_2 at O is the axis $y = 0$. We consider the expansions of y

$$\text{on } C_6\colon\ y = b_2x^2 + b_3x^3 + \cdots + b_kx^k + \cdots;$$
$$\text{on } C_2\colon\ y = c_2x^2 + c_3x^3 + \cdots + c_kx^k + \cdots,$$

which represent the curve C_6 and the conic C_2 respectively in the neighborhood of the origin. Let $A(x,y) = 0$ be the equation of the conic C_2. We may write then

$$(13)\quad \begin{cases} \psi_6(x,y) = (y - b_2x^2 - b_3x^3 - \cdots - b_kx^k - \cdots)\bar{\psi}(x,y); \\ A(x,y) = (y - c_2x^2 - c_3x^3 - \cdots - c_kx^k - \cdots)\bar{A}(x,y), \end{cases}$$

where $\bar{\psi}(x,y)$ and $\bar{A}(x,y)$ are polynomials in y, whose coefficients are functions of x, which are regular in the neighborhood of the value $x = 0$. Moreover, $\bar{\psi}(0,0) \neq 0$ and $A(0,0) \neq 0$.

We now have the following relation:

$$(14)\quad \begin{aligned} &(y - b_2x^2 - b_3x^3 - \cdots - b_kx^k - \cdots)^2[\bar{\psi}(x,y)]^2 - [\phi_4(x,y)]^3 \\ &= (y - c_2x^2 - c_3x^3 - \cdots - c_kx^k - \cdots)^2[\bar{A}(x,y)]^2 f(x,y). \end{aligned}$$

It should be noticed that $f(0,0) \neq 0$. Our proof will consist in showing that (14) cannot hold unless $a_k = b_k$ for every value of k, which is impossible, since this would mean that the curve C_6, and hence also C_4, contain the conic C_2 as a component, which contradicts the fact that C_6 and C_4 have no common components. Let us suppose that the first $k - 2$ coefficients $(k \geqq 2)$ of the above expansions are alike:

$$(15)\quad b_2 = c_2,\ b_3 = c_3, \cdots,\ b_{k-1} = c_{k-1}.$$

We propose to prove that $b_k = c_k$. Let

$$[\bar{\psi}(0,0)]^2 = \bar{\psi}_0^{\,2} \neq 0,\quad [\bar{A}(0,0)]^2 f(0,0) = a_0 \neq 0.$$

Recalling that $\phi_4(0,0) = 0$, we deduce immediately from (14)

$$(16)\quad \bar{\psi}_0^{\,2} = a_0.$$

If we put in (14)

$$(17)\quad y = y_1 = b_2x^2 + b_3x^3 + \cdots + b_{k-1}x^{k-1},$$

the functions $\bar{\psi}(x,y)$ and $\bar{A}(x,y)$ become integral power series in x, and we easily conclude, in view of (15), that the polynominal in x, $\phi_4(x,y_1)$, cannot

contain terms of degrees less than $2k/3$. If k is not divisible by 3, we put in (14)

$$y = y_2 = b_2x^2 + b_3x^3 + \cdots + b_{k-1}x^{k-1} + b_kx^k. \tag{17'}$$

We then observe that after this substitution is made the left-hand member of (14) does not involve terms of degree $\leqq 2k$ in x, and that in the right-hand member the coefficient of x^{2k} is $a_0(b_k - c_k)^2$. We deduce that $c_k = b_k$, which proves our assertion for the case $k \not\equiv 0$ (mod. 3).

Let now k be divisible by 3, $k = 3k_1$. If we again use the substitution (17) and if we denote by β_{2k_1} the coefficient of x^{2k_1} in $\phi_4(x, y_1)$, we obtain from (14) by equating the coefficients of x^{2k_1},

$$\psi_0^2 b_k^2 - \beta^3_{2k_1} = a_0c_k^2. \tag{18}$$

If $\beta_{2k_1} = 0$, then the reasoning employed above, in the case $k \not\equiv 0$ (mod. 3), can be used again in order to prove that $b_k = c_k$. We may then suppose that $\beta_{2k_1} \not\equiv 0$. We use the substitution (17′) and we first prove that as in $\phi_4(x, y_1)$ so also in $\phi_4(x, y_2)$ the term of lowest degree in x is $\beta_{2k_1}x^{2k_1}$. In fact, let $y = y_1'$, $y = y_2'$, etc., be the Puiseux expansions in the neighborhood of the origin of the different branches of the function y defined by the equation $\phi_4(x, y) = 0$ (y_1', y_2', etc., denoting fractional power series in x). Then we can write

$$\phi_4(x, y) = \rho(y - y_1')(y - y_2') \cdots,$$

where ρ is a constant or a polynominal in x. By the hypothesis made on the function $\phi_4(x, y_1)$ it follows that each of the series $y_1 - y_1'$, $y_1 - y_2' \cdots$ contains terms of lowest degree $\leqq 2k_1$. Since $y_2 = y_1 + b_kx^{3k_1}$, it is obvious that the terms of lowest degree of the series $y_1 - y_1'$, $y_1 - y_2', \cdots$, coincide with the terms of lowest degree of the corresponding series $y_2 - y_1'$, $y_2 - y_2', \cdots$, which proves that the terms of lowest degree of $\phi_4(x, y_1)$ and of $\phi_4(x, y_2)$ are the same.

If we now put in (14) $y = y_2$ and if we equate the coefficients of x^{2k_1}, we obtain

$$-\beta^3_{2k_1} = a_0(b_k - c_k)^2,$$

which combined with (16) and (18) yields

$$(b_k - c_k)^2 = c_k^2 - b_k^2. \tag{19}$$

In a similar way we obtain

$$(c_k - b_k)^2 = b_k^2 - c_k^2, \tag{19'}$$

by using the substitution $y = b_2x^2 + b_3x^3 + \cdots + b_{k-1}x^{k-1} + c_kx^k$.

From (19) and (19′) it follows that $b_k = c_k$. Q. E. D.

THE JOHNS HOPKINS UNIVERSITY.

CONSTRUCTION OF PENCILS OF EQUIANHARMONIC CUBICS.

By Jacob Yerushalmy.

1. *Introduction.* By Salmon's theorem, the cross-ratio of the four tangents, which may be drawn to a plane cubic curve from a point on it, is constant for the cubic. A certain function of the cross-ratio, which is rational in the coefficients of the cubic, constitutes the only absolute rational invariant of the cubic. In fact, if α is one of the values which this cross-ratio assumes, then

$$J = 4(\alpha^2 - \alpha + 1)^3/(\alpha + 1)^2(1 - 2\alpha)^2(2 - \alpha)^2$$

is the absolute rational invariant of the cubic. This invariant is known as the modulus of the cubic.

In terms of the two relative invariants S and T of the cubic, which are of degrees 4 and 6 respectively in the coefficients of the cubic,

$$J = S^3/T^2$$

and hence involves the coefficients to the 12th degree.

If $\alpha = -1$, then $J = \infty$, $T = 0$ and the cubic is said to be *harmonic.*

If $\alpha = -\epsilon(\epsilon^3 = 1)$, then $J = S = 0$ and the cubic is said to be *equianharmonic.*

If $\alpha = 1$, then $S^3 - T^2 = 0$ and the cubic has a double point.

If α is indetermined, $S = T = 0$ and the cubic has a cusp.

Since J involves the coefficients of the cubic to the 12th degree, there are in an arbitrary pencil of cubics 12 cubics of an assigned generic modulus, but only six harmonic and four equianharmonic cubics.

2. *Pencils of Cubics of Equal Modulus.* O. Chisini * proposes to determine all the pencils having the property that all the cubics of one pencil have the same modulus. He succeeds in determining them in the following manner.

He observes that to a pencil of cubics all having the same modulus corresponds a pencil of lines cutting a quartic curve in quadruples of points all having the same cross-ratio, and conversely. He then determines all the

* "Sui fasci di cubiche a modulo costante," *Rendiconti del Circolo Matematico di Palermo,* Vol. 41 (1916).

quartic curves admitting the above property and studying their singularities he arrives at the construction of the pencils of cubics.

In the equianharmonic case Chisini proves that every quartic that is cut by all the lines of a pencil in equianharmonic quadruples is represented by an equation of the type

$$f_4 = \frac{\partial^2 \phi_3(x_1x_2x_3)}{\partial x_3{}^2}\,\phi_3(x_1x_2x_3) - \tfrac{1}{2}\left[\frac{\partial \phi_3(x_1x_2x_3)}{\partial x_3}\right]^2 = 0,$$

where $\phi_3 = 0$ is an arbitrary cubic not passing thru the center of the pencil of lines which is taken to be the point $C \equiv (0, 0, 1)$. It is easily verified that the cubic $\phi_3 = 0$ is the polar of the point C with respect to the quartic $f_4 = 0$. He also shows that the six points C_1, C_2, C_3, C_4, C_5, C_6 of intersection of $\phi_3 = 0$ with the polar conic of C with respect to $\phi_3 = 0$ are flex-points for $f_4 = 0$ whose flex-tangents pass thru C. Obviously the cubic $\phi_3 = 0$ touches $f_4 = 0$ at the six points C_i.

Chisini, however, does not give the actual construction of the pencils of equianharmonic cubics. He only points out that such pencils are characterized by possessing six cuspidal cubics (corresponding to the six flex-tangents CC_i of the above quartic). It is the object of this paper to make a closer investigation of these pencils of cubics with special reference to their actual construction. We show that every such pencil is contained in a net of equianharmonic cubics thru six base-points. These base-points form the vertices of two in-circumscribed triangles of an arbitrary cubic curve, which are three-fold perspective from the vertices of a third in-circumscribed triangle of the same cubic.

3. *Cremona Transformations Leaving a Pencil of Equianharmonic Cubics Invariant.* It is well known that a general elliptic cubic is transformed into itself by 8 cyclic collineations of period three whose fixed points do not lie on the cubic. These collineations are given in terms of the abelian parameter u in the form

$$u' = u + \omega/3 \qquad (\omega \text{ is a period}).$$

In addition to these general transformations, an equianharmonic cubic is transformed into itself by singular homographies and homologies of period three with three fixed points on the cubic. In terms of the abelian parameter u these singular transformations are given by

$$u' = + \epsilon u + b, \qquad (\epsilon^3 = 1).$$

The existence of such homographies and homologies characterizes the equianharmonic cubics.

If, then, we have a linear system of equianharmonic cubics, say a pencil, the homology may be fixed or variable, i. e., there may exist one homology leaving invariant all the cubics of the pencil, or the homology will vary from cubic to cubic. However Chisini proves (*loc. cit.*, p. 90) that if the homology is fixed the pencil is special, and on a general pencil the homology is variable. It is natural, therefore, to look for Cremona Transformations of the plane into itself leaving the cubics of a pencil of equianharmonic cubics invariant. For this purpose we must investigate more closely the quartic curve $f_4 = 0$ which is cut by a pencil of lines $\{C\}$ in equianharmonic quadruples of points.

As is known, a cubic surface F_3 may be mapped on a double plane with a quartic branch-curve by projection from a point O on the surface. Consider, therefore, the plane π of $f_4 = 0$ as the projection of a cubic surface F_3, and $f_4 = 0$ as the branch-curve. The equation of $f_4 = 0$ is, as we noted,

$$f_4 \equiv \frac{\partial^2 \phi_3}{\partial x_3{}^2} \phi_3 - \tfrac{1}{2} \left(\frac{\partial \phi_3}{\partial x_3} \right)^2 = 0$$

and $\phi_3 = 0$ is the polar of C (the center of $\{C\}$) with respect to $f_4 = 0$. Take an arbitrary line a of $\{C\}$. It cuts $f_4 = 0$ in an equianharmonic quadruple of points E_1, E_2, E_3, E_4 and $\phi_3 = 0$ in three points P_1, P_2, P_3 which constitute the polar group of C with respect to $E_1, \cdots E_4$. To a will correspond on F_3 a plane equianharmonic cubic curve Ψ_3 cut out by the plane thru a and O (O is the center of projection). To P_1, P_2, P_3 will correspond on Ψ_3 two triples of points. We proceed to prove that each triple of points is a group of three fixed points of a singular birational transformation T_3, cyclic of order 3, of the cubic Ψ_3 into itself. What we have to prove is, in fact, the following:

THEOREM. *The lines joining any point O on an equianharmonic cubic to the ∞^1 groups of three fixed points, on it, of the ∞^1 singular birational transformations cyclic of period 3, of the cubic into itself form a $g_3{}^1$ in the pencil $\{O\}$, and this $g_3{}^1$ is precisely the $g_3{}^1$ obtained by taking the polar group of a variable line of the pencil with respect to the four tangents to the cubic from O.*

Proof. Let Ψ_3 be an equianharmonic cubic and O a point on it. Let γ be a transformation of Ψ_3 into itself having three fixed points A_1', A_2', A_3' on the cubic. Let ω be the transformation determined by the $g_2{}^1$ cut out on Ψ_3 by the pencil $\{O\}$, i. e. the transformation which interchanges the points of each pair of the $g_2{}^1$.

Evidently $\omega^{-1}\gamma\omega$ is a transformation of Ψ_3 into itself having as fixed

6

points the three further intersections A_1'', A_2'', A_3'' of the lines OA_1', OA_2', OA_3' with Ψ_3. Hence given any line of the pencil, each of the two points in which it cuts Ψ_3 outside of O will define, by a known result, two other points which together with it form a group of three fixed points for some birational transformation of Ψ_3 into itself, and by the previous remark these two triples will form three groups of the $g_2{}^1$ mentioned above. It follows that the series ∞^1 of triples of lines mentioned in the theorem is such that each line belongs to one and only one triple, and therefore this series is a $g_3{}^1$.

To prove that this $g_3{}^1$ coincides with the $g_3{}^1$ of the polar groups we take the trace of $\{O\}$ on a line l.

Consider the transformation γ of Ψ_3 into itself having one of its fixed points at the point of contact C of one of the tangents thru O. γ will leave invariant the $g_2{}^1$ cut out by $\{O\}$ since it leaves one group of it (the point C counted twice), invariant, and will permute cyclically the three points of contact of the other 3 tangents to Ψ_3 from O. The other two fixed points of γ are on a line with O. We have in $\{O\}$ a cyclic projectively of order 3. The trace of it on l may be given by $x' = \epsilon x$ $(\epsilon^3 = 1)$, having the points 0 and ∞ as invariant points. The equianharmonic quadruple of points on l can be taken to be 0, 1, ϵ, ϵ^2 and the polar group of any point (a_1, a_2) with respect to it is given by

$$(1) \qquad 4a_1x_1^3 - 3a_2x_1x_2^2 - a_1x_2^3 = 0 \quad \text{or} \quad 4ax^3 - 3x - a = 0.$$

The polar group of the trace of OC on l is the polar group of 0 with respect to 0, 1, ϵ, ϵ^2. Putting in (1) $a = a_1/a_2 = 0$ we obtain the point O and the point ∞ counted twice, and this triple of points obviously coincides with the traces on l of the lines joining O to the three fixed point of γ. The same will hold for any of the 4 groups having one of the 4 points of contact as a fixed point, hence the $2g_3{}^1$'s coincide. q. e. d.

From this theorem we conclude that the two triples of points on F_3 are each a group of fixed points of some birational transformation of Ψ_3 into itself.

To ϕ_3 will correspond on F_3 a curve of order 9, C_9. It will be seen later from the analytical expression for F_3 that C_9 is degenerate, but it can also be seen from the following consideration. Since $\phi_3 = 0$ touches the branch curve $f_4 = 0$ whenever they meet, the doubly-covered cubic $\phi_3 = 0$, i. e. C_9 does not possess branch-points. Moreover the six points of intersection of $\phi_3 = 0$ and $f_4 = 0$ lie on a conic $(\Psi_2 = \partial\phi_3/\partial x_3 = 0)$ and by a known theorem * C_9 is reducible. It will be seen later that C_9 breaks up into a plane

* See for instance F. Enriques and O. Chisini, *Courbes et Fonctions Algébriques d'une Variable*, Chap. IV, p. 444.

cubic curve which may be supposed to coincide with the cubic $\phi_3=0$ itself and another sextic curve. The two triples of points on F_3 corresponding to P_1, P_2, P_3 are, hence, separable. One group of 3 points will be on $\phi_3=0$ and the other group will be on the sextic.

Consider only the cubic $\phi_3=0$. It is traced out by the groups of fixed points of the singular birational transformations, cyclic of order 3, sending into themselves the cubics of the pencil which we obtain on F_3 as the line a varies in the pencil $\{C\}$, i. e., the cubics cut out by the planes thru OC.

These singular transformations of the cubics of the pencil are, in fact, homologies because the three fixed points on each cubic are on a line (the line of the pencil $\{C\}$ corresponding to the cubic).

Since thru each point of F_3 there passes only one cubic of the pencil we have a Cremona transformation Γ of the space sending F_3 into itself defined in the following way: For any point P take the plane of the pencil containing it. In this plane we have an homology sending P into P' in the same plane. P' is the homologous point of P under Γ. Evidently Γ possesses on F_3 an entire curve of invariant points, the curve $\phi_3=0$ of fixed-points.

4. *The Equation of F_3.* Consider the equation

$$\frac{\partial^2\phi_3(x_1x_2x_3)}{\partial x_3{}^2}x_4{}^2+(2)^{1/2}\frac{\partial\phi_3(x_1x_2x_3)}{\partial x_3}x_4+\phi_3(x_1x_2x_3)=0.$$

It is an equation of a cubic surface F. If we project F from the point $O\equiv(0,0,0,1)$ on it upon the plane $x_4=0$ we get as branch-curve

$$(\partial^2\phi_3/\partial x_3{}^2)\phi_3-\tfrac{1}{2}(\partial\phi_3/\partial x_3)^2=0$$

which coincides with our $f_4=0$.

The following consideration shows that we can really take the above equation to represent our F_3.

The plane $x_4=0$ is a double plane both for F and the plane α of the pencil of equianharmonic cubics (see Chisini, *loc. cit.*, p. 62), the branch-curve being the same in both cases. The net of cubics on F cut out by the bundle of planes thru O will be mapped into the net of cubics thru the seven base-points in α. The system ∞^3 of the plane sections of F will go by this mapping into the web of cubics thru 6 of the 7 base-points, the seventh base-point will correspond to O. This web will therefore contain the pencil of equianharmonic cubics and we can take as the equation of F_3:

$$F_3\equiv(\partial^2\phi_3/\partial x_3{}^2)x_4{}^2+(2)^{1/2}(\partial\phi_3/\partial x_3)x_4+\phi_3=0.$$

From this equation it is evident that the plane $x_4=0$ cuts out on F_3

the cubic $\phi_3 = 0$, and that the C_9 on F_3 corresponding to $\phi_3 = 0$ breaks up into $\phi_3 = 0$ itself and another sextic curve. $\phi_3 = 0$ is, therefore, the curve of invariant points for the Cremona Transformation Γ sending F_3 into itself.

5. *Every Pencil of Equianharmonic Cubics is Contained in a Net of Such Cubics.* The only points on F_3 which may be fundamental points for Γ are the three points O, O_1, O_2, in which OC meets F_3. We proceed to prove, however, that they are ordinary points and consequently Γ possesses no fundamental points on F_3.

The singular birational transformations sending the cubics of the pencil, cut out on F_3 by the planes on OC, into themselves are, as we have noted, homologies. Therefore the three fixed points are flexes for the cubic, and it will be shown later that the line joining them is a side of the Hessian triangle of the cubic. The pencil $\{C\}$ is, hence, composed of the flex lines for the corresponding cubics. Each line forms the axis of the homology; the center being the point common to the three flex tangents at the invariant points.

Since O, O_1, O_2 are not on $\phi_3 = 0$, one of the following two cases may happen: either these points go by each homology into different points on the different cubics and will, therefore, be fundamental points for Γ; or they form for each homology a cycle. This will happen if for each cubic the three flex tangents at the invariant points meet on OC. In this case O, O_1, O_2 are ordinary points for Γ which possesses, then, no fundamental points on F_3. To prove that this latter case takes place, it is sufficient, remembering that on a P_1, P_2, P_3 are the polar group of C with respect to E_1, E_2, E_3, E_4, to prove the following

THEOREM. *The lines joining any point O of an equianharmonic cubic to any three flexes P_1, P_2, P_3 on a side of the Hessian triangle form the polar group of the line joining O to the point K common to the three flex-tangents at P_1, P_2, P_3 with respect to the 4 tangents drawn from O to the cubic.*

Proof. Let the equation of the equianharmonic cubic be

$$\Psi_3 \equiv x_1^3 + x_2^3 + x_3^3 = 0.$$

The Hessian of $\Psi_3 = 0$ is the cubic $x_1x_2x_3 = 0$ composed of the three lines $x_1 = 0$, $x_2 = 0$, $x_3 = 0$.

Let $O \equiv (y_1, y_2, y_3)$ be any point on $\Psi_3 = 0$.

The four tangents from O to $\Psi_3 = 0$ are given by

$$3(y_1x_1^2 + y_2x_2^2 + y_3x_3^2)^2 - 4(x_1^3 + x_2^3 + x_3^3)(y_1^2x_1 + y_2^2x_2 + y_3^2x_3) = 0.$$

Let P_1, P_2, P_3 be the three flexes on the line $x_2 = 0$. The point K in this case will be the point $x_1 = x_2 = 0$ which is the opposite vertex of the flex triangle. Let the pencil of lines $\{O\}$ be projected upon $x_2 = 0$. The three flexes P_1, P_2, P_3 are given by

$$x_1^3 + x_3^3 = 0.$$

The trace of the line OK is the point (y_1, y_3). The equianharmonic quadruple is given by

$$\phi_4 = (y_1 x_1^2 + y_3 x_3^2)^2 - 4(x_1^3 + x_3^3)(y_1^2 x_1 + y_3^2 x_3) = 0.$$

The polar triple of (y_1, y_3) with respect to $\phi_4 = 0$ is given by

$$y_1 \partial\phi_4/\partial x_1 + y_3 \partial\phi_4/\partial x_3 = 0$$

or

$$\begin{aligned} & y_1\{12x_1y_1(y_1x_1^2 + y_3x_3^2) - 4y_1^2(x_1^3 + x_3^3) - 12x_1^2(y_1^2x_1 + y_3^2x_3)\} \\ & \quad + y_3\{12x_3y_3(y_1x_1^2 + y_3x_3^2) - 4y_3^2(x_1^3 + x_3^3) - 12x_3^2(y_1^2x_1 + y_3^2x_3)\} = 0 \\ & = 12(y_1x_1^2 + y_3x_3^2)(y_1^2x_1 + y_3^2x_3) - 4(x_1^3 + x_3^3)(y_1^3 + y_3^3) \\ & \quad - 12(y_1x_1^2 + y_3x_3^2)(y_1^2x_1 + y_3^2x_3) = 0 \\ & = x_1^3 + x_3^3 = 0, \end{aligned}$$

which coincides with the three flexes. q. e. d.

We conclude that the three points O, O_1, O_2 form a cycle for each homology, and they are ordinary points for Γ. Moreover, since the point K in each cubic must form with the three points O, O_1, O_2 an equianharmonic quadruple, it is the same point for all the cubics of the pencil. The transformation Γ has therefore the point K as an invariant point. Also every point of the plane $x_4 = 0$ is invariant for Γ, since we have in this plane the pencil of lines $\{C\}$ each of which is the axis of an homology. We also see that the projectivity involved on the line OC by the homology on a generic plane on OC does not vary as the plane varies in the pencil. Hence the transformation Γ does not possess fundamental points, and consequently is a collineation. More precisely, Γ is an homology in space, cyclic of order three, with the point K as center and $x_4 = 0$ as plane of invariant points.

The net of cubics cut out on F_3 by the bundle of planes on K is a net of invariant cubics for Γ since the planes of the bundle are invariant. Each cubic of the net has three invariant points on it (the three points in which it cuts the plane $x_4 = 0$) and is therefore equianharmonic.

By mapping the cubic surface F_3 on the plane α of the original pencil of equianharmonic cubics we conclude that every pencil of equianharmonic cubics is contained in a net of such cubics thru 6 of the 9 base points of the pencil.

6. *A Canonical Form for F_3 and the Equation of the Homology.* Let us determine the coördinates of the point K. If we write $\phi_3 = 0$ in the canonical form

$$\phi_3 \equiv x_1^3 + x_2^3 + x_3^3 + 6\lambda x_1x_2x_3 = 0,$$

F_3 will take the form

$$F_3 \equiv x_1^3 + x_2^3 + x_3^3 + 6x_3x_4^2 + 3(2)^{1/2}x_3^2x_4 + 6(2)^{1/2}\lambda x_1x_2x_4 + 6\lambda x_1x_2x_3 = 0.$$

Consider the cubic Ψ_3 cut out on F_3 by the plane $x_1 = \mu x_2$:

$$\Psi_3 \equiv (1 + \mu^3)x_2^3 + x_3^3 + 6x_3x_4^2 + 3(2)^{1/2}x_3^2x_4 + 6(2)^{1/2}\lambda\mu x_2^2x_4 + 6\lambda\mu x_2^2x_3 = 0.$$

This is an equianharmonic cubic and its Hessian is composed of a flex triangle. The equation of the Hessian is found to be

$$H \equiv x_4\{[(1 + \mu^3)x_2 + 2\lambda\mu(x_3 + (2)^{1/2}x_4)](x_3 + (2)^{1/2}x_4) - 2\lambda^2\mu^2x_2^2\} = 0.$$

The second factor breaks up into the product of two linear factors by completing the square and taking the difference of the two squares, and we obtain the Hessian triangle. Since $x_4 = 0$ is one of the sides we conclude that each of the lines of the pencil $\{C\}$ is a side of the Hessian triangle of the cubic corresponding to it on F_3.

To determine the coördinates of K, we consider the cubic cut out on F_3 by the particular plane thru $x_1 = 0$ and O. The Hessian triangle of this cubic is found to be

$$x_4x_2[(2)^{1/2}x_3 + 2x_4] = 0.$$

The sides of the triangle are the lines

$$x_2 = 0; \quad x_4 = 0; \quad (2)^{1/2}x_3 + 2x_4 = 0.$$

The three tangents at the flexes on $x_4 = 0$ meet on the opposite vertex, which is the point of intersection of $x_2 = 0$ and $(2)^{1/2}x_3 + 2x_4 = 0$, or in the point $K \equiv [0, 0, 1, -(2)^{1/2}/2]$.

For simplicity send the point K to $(0, 0, 1, 0)$ and interchange the planes $x_3 = 0$ and $x_4 = 0$.

The transformation sending the points

$$(0, 0, 0, 1),\ (0, 0, -(2)^{1/2}/2, 1),\ (0, 0, 1, 0) \text{ into } (0, 0, 0, 1),\ (0, 0, 1, 0),\ (0, 0, \epsilon, 1)$$

is found to be $x_1 = x_1'$; $x_2 = x_2'$; $x_3 = \epsilon x_3'$; $x_4 = (2)^{1/2}(x_4' - \epsilon x_3')$.

Applying this transformation to F_3 after having interchanged x_3 and x_4 the equation of F_3 turns out to be

$$F_3 \equiv x_1^3 + x_2^3 - 2(2)^{1/2}x_3^3 + 2(2)^{1/2}x_4^3 + 6(2)^{1/2}\lambda x_1x_2x_4 = 0,$$

and the homology in space sending F_3 into itself is evidently

$$\Gamma; \quad x_1' = x_1, \quad x_2' = x_2, \quad x_3' = \epsilon x_3, \quad x_4' = x_4.$$

7. *A Net of Equianharmonic Cubics thru Six Base-Points.* In order to obtain the equianharmonic net of cubics in the plane we map F_3 on a plane α.

As is known, a cubic surface may be mapped on a plane by means of a system ∞^3 of cubics thru six base-points in the plane, so that the cubics of this system correspond to the plane sections of the cubic surface. The correspondence between the points of the plane and the points of the surface is $(1,1)$ except for the six base-points i $(i = 1, 2, \cdots 6)$ to which correspond on the surface six lines denoted by a_i. Also to the six conics thru 5 of the base-points correspond on the cubic surface six lines b_i (the line b_i corresponding to the conic leaving out the point i). Finally to the 15 lines joining two of the base points i and j correspond on the surface the lines c_{ij}. In all we have 27 lines on the cubic surface.

F_3 is a particular cubic surface since there are collineations sending it into itself, and when mapped on α the six base points will be in particular position which we proceed to determine.

Evidently to Γ will correspond in α a transformation T. For a point P on F_3 is mapped into a point Q in α. Γ sends P into a point P' which is mapped into a point Q' in α. The correspondence between Q and Q' is $(1,1)$ and algebraic, and is therefore a Cremona transformation in the plane.

Since Γ has no fundamental points on F_3, the Cremona transformation T will have no fundamental points in α outside the six base-points. Let us determine the order of T, i. e., the order of the curves into which T transforms the lines of α.

Consider an arbitrary line a of α. It is mapped into a twisted cubic C on F_3. Γ sends C into another twisted cubic $\bar{C}$ which is mapped into a curve S of α. As is known there are seventy two systems ∞^2 of twisted cubics on F_3. One system is mapped into the lines of α. Twenty systems are mapped into conics thru three of the base-points, thirty systems into cubics with a double point at one of the base-points and passing simply thru four of the remaining five base-paints. Twenty systems go into quartics with double points in three of the base-points and simple points at the other three base-points, and finally one system is mapped into quintics having double-points at the six base-points. To determine into what system Γ transforms the system of cubics corresponding to the lines of α we must find the number of intersections of C and $\bar{C}$.

C cuts $x_3 = 0$ in three points which are invariant for Γ and belong, hence, also to $\bar{C}$. If C and $\bar{C}$ have any more common points they must be such points on C which are the transform of points on C. A point and its transform by Γ must be on a line with K. However, there is only one double-secant of C thru K, and since Γ is cyclic of order three, one of the two points in which the double-secant cuts C must go by it into the other point. Hence, one of these two points is common to C and $\bar{C}$. These two curves intersect therefore in four points, and so do S and a in α. Hence S is a quartic curve passing simply thru three of the base-points, say 1, 2, 3 and doubly thru the other base-points 4, 5, 6. The transformation is biquadratic and the fundamental points of the inverse transformation (T^{-1}) coincide with those of the direct transformation. The homoloidal net Σ' of the inverse transformation being known

$$\Sigma' : \quad (1^1, 2^1, 3^1, 4^2, 5^2, 6^2)_4,$$

it is easy to determine the homoloidal net Σ of T. T being cyclic of order three must transform the system Σ' into Σ. A quartic C_4 of Σ must therefore have four variable intersections with a quartic C_4' of Σ'. Let Σ have multiplicities r_1, r_2, r_3 at 1, 2, 3 and s_1, s_2, s_3 at 4, 5, 6. We have

$$r_1 + r_2 + r_3 + 2(s_1 + s_2 + s_3) = 16 - 4 = 12$$

and since three of them must have the value two and the other three must have the value one, we have the result

$$r_1 = r_2 = r_3 = 2, \qquad s_1 = s_2 = s_3 = 1,$$

and hence

$$\Sigma : \quad (1^2, 2^2, 3^2, 4^1, 5^1, 6^1)_4.$$

Corresponding to the cubic $\phi_3 = 0$ on F_3, we have in the system ∞^3 of cubics in α a cubic ϕ each point of which is invariant for T. Consider the base-point 4. It is a simple fundamental point and goes by T into a line l which has no variable intersections with the quartics of Σ' and is therefore on two of the three points 4, 5, 6. This says that on F_3, a_4 goes into c_{ij} $(ij = 4, 5, 6)$ by Γ. But c_{ij} must pass thru the point of intersection of a_4 and $\phi_3 = 0$ and therefore the line l in α passes thru 4. It may be either $(45)_1$ or $(46)_1$. Let $T(4) = (45)_1$, then $T(5) = (56)_1$, $T(6) = (64)_1$. Moreover, since $\phi_3 = 0$ and c_{45} have a point in common on a_4, the line $(45)_1$ and the cubic ϕ have at 4 the same direction. $(45)_1$ is, hence, tangent to ϕ at 4. In the same way it is seen that $(56)_1$ and $(64)_1$ are tangent to ϕ at 5 and 6 respectively. The three points 4, 5, 6 form therefore an in-circum-

scribed triangle in the cubic ϕ. Applying the same reasoning for T^{-1} we see that the other three fundamental points 1, 2, 3 form the vertices of another in-circumscribed triangle in the same cubic.

The point 1 goes by T into a conic passing thru 1. This conic goes thru 4, 5, 6 since it has no variable intersections with the quartics of Σ'. 1 is transformed, hence, into either $(12456)_2$ or $(13456)_2$. Let $T(1)=(12456)_2$ then $T(2)=(23456)_2$ and $T(3)=(31456)_2$. We have for T

$$\begin{array}{l}1 \to (12456)_2 \to (13)_1 \to 1\\ 2 \to (23456)_2 \to (12)_1 \to 2\\ 3 \to (31456)_2 \to (23)_1 \to 3\\ 4 \to (45)_1 \to (12346)_2 \to 4\\ 5 \to (56)_1 \to (12345)_2 \to 5\\ 6 \to (64)_1 \to (12356)_2 \to 6.\end{array}$$

In order to obtain the complete configuration of the six base-points, consider for example the point 1. We have by T

$$1 \to (12456)_2 \to (13)_1 \to 1.$$

On F_3 we have the cycle $(a_1\, b_3\, c_{31})$. These three lines must concur at a point on $\phi_3=0$ (the invariant point on a_1), and hence both the conic (12456) and the line (31) are tangent to ϕ at 1. The same is true for the other five conics. The complete configuration of the six base-points is therefore as follows: (1, 2, 3) (4, 5, 6) are the vertices of two in-circumscribed triangles on ϕ, such that the lines (13), (12), (23), (45), (56), (64) and the conics (12456), (23456), (31456), (12346), (12345), (12356) are tangent to ϕ at the points 1, 2, 3, 4, 5, 6 respectively.

As is known there are twenty-four in-circumscribed triangles in an arbitrary cubic ϕ. For, if (4, 5, 6) be such a triangle, and u the abelian parameter of the cubic chosen so that for three points on a line the sum of the values of this parameter should be equal to zero (mod. periods), then we have

$$\begin{array}{l}2u_4+u_5\equiv 0 \ (\text{mod. } \omega, \omega')\\ 2u_5+u_6\equiv 0 \ (\text{mod. } \omega, \omega')\\ 2u_6+u_4\equiv 0 \ (\text{mod. } \omega, \omega').\end{array}$$

From which

$$u_4\equiv(\alpha\omega+\alpha'\omega')/9.$$

Giving to α and α' all the values from 0 to 8 we get 81 points from which we have to exclude the 9 flexes, leaving 72 points forming 24 triangles.

If we start with any ninth of a period for u_4, then u_5 and u_6, and hence

the triangle (4, 5, 6), are determined. In order to determine the second triangle (1, 2, 3) we must put the condition that $(12346)_2$ should touch ϕ at 4. Hence

$$(2) \qquad 2u_4 + u_5 + u_1 + u_2 + u_3 \equiv 0 \pmod{\omega, \omega'}.$$

But since (1, 2, 3) is also an in-circumscribed triangle, the following relations exist

$$u_3 \equiv -2u_1 \pmod{\omega, \omega'}$$
$$u_2 \equiv 4u_1 \pmod{\omega, \omega'}.$$

Substituting in (2) we have

$$2u_4 + u_1 \equiv (\beta_\omega + \beta'_{\omega'})/3,$$
$$u_1 \equiv -2u_4 + 1/3 \text{ of a period} \equiv u_4 + 1/3 \text{ of a period}.$$

We have 9 thirds of a period, and it would seem that associated with any triangle there are 8 more each of which together with it will give the required six points. However, finding the values of u_2 and u_3

$$u_2 \equiv 4u_4 + 1/3 \text{ of a period} \equiv u_4 + 1/3 \text{ of a period}$$
$$u_3 \equiv -8u_4 + 1/3 \text{ of a period} \equiv u_4 + 1/3 \text{ of a period}$$

we see that these also are obtained by adding to u_4 a third of a period, and hence the nine thirds of a period give rise to only 3 triangles. One triangle is to be excluded; this is the one obtained by adding to u_4 the third periods 0, $3u_4$, $6u_4$, which give the triangle (4, 5, 6) over again. Hence, starting with any one point (a ninth of a period), the triangle containing it is determined and with it there are determined two associated triangles, each of which can be taken with the original one to form the base of the net.

In fact, start with a ninth of a period $u_1 \equiv \omega/9$. The three triangles are

$$\begin{array}{lll} u_1, & u_2 \equiv u_1 + \omega/3, & u_3 \equiv u_1 + 2\omega/3; \\ u_4 \equiv u_1 + \omega'/3, & u_5 \equiv u_1 + 2\omega/3 + \omega'/3, & u_6 \equiv u_1 + \omega/3 + \omega'/3; \\ u_7 \equiv u_1 + 2\omega'/3, & u_8 \equiv u_1 + 2\omega/3 + 2\omega'/3, & u_9 \equiv u_1 + \omega/3 + 2\omega'/3. \end{array}$$

It is easily verified from the above that any two of the three triangles satisfy our conditions, and any two of them can be taken to form the base.

It is also easily seen that any two of the above three triangles are threefold perspective from the vertices of the third. For

$$u_1 + u_4 + u_8 \equiv u_2 + u_5 + u_8 \equiv u_3 + u_6 + u_8 \equiv 0;$$
$$u_1 + u_5 + u_7 \equiv u_2 + u_6 + u_7 \equiv u_3 + u_4 + u_7 \equiv 0;$$
$$u_1 + u_6 + u_9 \equiv u_2 + u_4 + u_9 \equiv u_3 + u_5 + u_9 \equiv 0.$$

Finally we want to show that if we take on an arbitrary cubic ϕ six points forming the vertices of two in-circumscribed triangles and satisfying the above conditions, then in the system ∞^3 of cubics thru these points there is contained a net of equianharmonic cubics.

Consider a biquadratic transformation T having these six points as fundamental points $(1^2, 2^2, 3^2, 4^1, 5^1, 6^1)$. The fundamental points of the inverse transformation will generally be some other six points $(1_0{}^1, 2_0{}^1, 3_0{}^1, 4_0{}^2, 5_0{}^2, 6_0{}^2)$ of the plane. T sends the cubics thru $(1, 2, \cdots 6)$ into the cubics thru $(1_0, 2_0, \cdots 6_0)$. It will send ϕ into some cubic ϕ_0.

The quartics of Σ' having double-points at 1, 2, 3 and passing simply thru 4, 5, 6 cut the cubic ϕ in ∞^2 groups of 3 points. But since

$$2(u_1 + u_2 + u_3) \equiv 2\omega/3,$$
$$u_4 + u_5 + u_6 \equiv \omega/3,$$

and hence

$$2(u_1 + u_2 + u_3) + u_4 + u_5 + u_6 \equiv 0,$$

it follows that the sum of the Abelian parameters at each group of three points is equal to zero, i. e. the three points of each group are on a line. Σ' cuts, therefore, out on the cubic ϕ the same $g_3{}^2$ cut out on it by the lines of the plane. The transformation T involves, therefore, a collineation γ sending ϕ into ϕ_0. γ sends the points $1, 2, \cdots 6$ into the points $1_0, 2_0 \cdots 6_0$. To prove this take for example the line (13). T sends it into one of the points 1_0, 2_0, 3_0 say into 1_0. Consider the pencil of quartics in Σ' degenerating into the fixed line (13) and the pencil of cubics $(2^2, 4^1, 5^1, 6^1, 1^1, 3^1)$. Then $g_3{}^1$ cut out by it on ϕ has a fixed point which falls at 1 since (13) is tangent to the cubic ϕ at this point. To it will correspond the $g_3{}^1$ cut out on ϕ_0 by the lines on 1_0, and the fixed point 1_0 corresponds to the fixed point 1. And by the same reasoning γ send 2 into 2_0 etc.

Multiplying T by γ^{-1} we get a Cremona transformation R having the fundamental points of both the direct and inverse transformations coincident. This transformation will leave invariant the linear system ∞^3 of cubics on $1, \cdots 6$. Moreover, each point of the cubic ϕ is an invariant point for R.

The cube of this transformation is a collineation in the plane. Because the lines of the plane go by it into the quartics of Σ': $(1^1, 2^1, 3^1, 4^2, 5^2, 6^2)$, which go by the same transformation into quartics having four variable intersections with those of Σ'. R sends, therefore, Σ' into Σ. Applying R to Σ we get again the lines of the plane. But a collineation having a cubic of invariant points is the identity, hence $R^3 = I$ and the transformation is cyclic of period 3.

If we consider then the linear space S_3 whose elements are the cubics of our system ∞^3, our transformation R will be a cyclic projectivity because it is an algebraic $(1, 1)$ correspondence between the elements. This projectivity has one invariant point (the cubic ϕ) and also every line on this point is invariant: in fact every pencil of cubics obtained by combining linearly any cubic of the system with ϕ goes into itself since the three base-points of the pencil outside the fundamental points are invariant. By duality, since the projectivity has a bundle of invariant lines, it must have a plane of invariant points. The transformation R in the plane has, therefore, a net of invariant cubics. Each of the cubics of this net has three fixed points on it (the three points in which it cuts the cubic ϕ) and is therefore equianharmonic.

It is of interest to note that the cubic ϕ is the locus of the cusps of all the cuspidal cubics of the net, because the cusp must be an invariant element and is therefore on the invariant cubic.

THE PLANAR IMPRIMITIVE GROUP OF ORDER 216.

By John Rogers Musselman.

Introduction.

In a paper,* which discussed the imprimitive group of order 5184 in S_3, the existence of an imprimitive group of order 216 in S_2 connected with the Hesse configurations was indicated. The purpose of this paper is to discuss this planar group and the geometry associated with it. Of special interest is a configuration of twelve triangles which possesses important properties. These are dealt with in Section II. In Section I is a study of four-fold perspective triangles. A new canonical form is used which leads to a ruler construction for them and to a compass construction for a Clebsch six-point.

I. Four-fold Perspective Triangles.

Four-fold perspective triangles have been discussed by Schröter,† Hess,‡ Valyi § and others—all using the same canonical form. As a knowledge of these triangles is necessary for the development of the next section of this paper it has seemed best to study them using a new canonical form, which has certain advantages over the former one. Among others it leads to a simple ruler construction for ordinary four-fold perspection, and to a compass construction for a Clebsch six-point.

Let us call the vertices of one triangle A, B, C; those of the second triangle a, b, c. We shall choose for our canonical form the coördinates of the six vertices to be

$$\begin{array}{lll} A:\ (1,0,0) & B:\ (0,1,0) & C:\ (0,0,1), \\ a:\ (1,1,1) & b:\ (r_1,r_2,r_3) & c:\ (s_1,s_2,s_3). \end{array}$$

Let us indicate the four-fold perspections as $A_aB_bC_c$, $A_bB_cC_a$, $A_cB_aC_b$, $A_aB_cC_b$. To be perspective in these four ways requires the following conditions on the six points

$$r_1s_2 = r_3s_1, \quad r_2s_3 = r_1s_2, \quad r_3s_1 = r_2s_3, \quad r_1s_3 = r_2s_1. \tag{1}$$

---

* Musselman, *American Journal of Mathematics*, Vol. 49 (1927), pp. 355-366.

† Schröter, *Mathematische Annalen*, Vol. 2 (1869), p. 553.

‡ Hess, *Mathematische Annalen*, Vol. 28 (1887), p. 167.

§ Valyi, *Archiv der Mathematik und Physik*, Vol. 70 (1884); Vol. A1 (1885), p. 320; *Monatshefte für Mathematik und Physik* (1898), p. 169.

As is well known the third condition above is merely a consequence of the first two and it is only three conditions on two triangles to be four-fold perspective. If we eliminate r from the three actual conditions we have $s_1(s_1s_2 - s_3^2) = 0$; if we eliminate s from the conditions we have $r_1(r_1r_3 - r_2^2) = 0$. Now $s_1 = 0$ together with conditions (1) would compel one of the points of the second triangle to coincide with one of the first triangle and we should not have two distinct triangles. Similarly if $r_1 = 0$. Hence we see that for four-fold perspective triangles point c must lie on the conic $x_1x_2 - x_3^2 = 0$ and point b on the conic $x_1x_3 - x_2^2 = 0$. The first conic is tangent to $\overline{AC}$ at A, tangent to $\overline{BC}$ at B and passes through a. The second conic is tangent to $\overline{AB}$ at A, tangent to $\overline{BC}$ at C and passes through a.

The equation of the line joining points b and c is

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ r_1 & r_2 & r_3 \\ s_1 & s_2 & s_3 \end{vmatrix} = 0,$$

which can be written, due to (1), as

$$r_1/s_1(s_2 - s_3)[-(s_2 + s_3)x_1 + s_1(x_2 + x_3)] = 0.$$

Now $s_2 \neq s_3$ else point c would coincide with point a, so the equation of the line is simply

$$(s_2 + s_3)x_1 - s_1x_2 - s_1x_3 = 0. \tag{2}$$

If we let points b and c run over the conics on which they lie and ask for the envelope of the line $\overline{bc}$ we find that all these lines pass through the point $(0, -1, 1)$. To locate this point p, note that if we call the point of intersection of $\overline{BC}$ and $\overline{Aa}$ by p', then p lies on $\overline{BC}$ and is the fourth harmonic of p' as to B and C. Moreover point p and line $\overline{Aa}$ set up the reflexion $\rho x_1' = x_1$, $\rho x_2' = x_3$, $\rho x_3' = x_2$ which interchanges the two conics $x_1x_2 - x_3^2 = 0$ and $x_1x_3 - x_2^2 = 0$.*

The above furnishes a construction for four-fold perspective triangles given one triangle and one vertex of the second triangle. Call the given triangle ABC and the fourth point a. Produce $\overline{Aa}$ to cut $\overline{BC}$ at p'. Construct p as the fourth harmonic of p' as to B and C. Construct any point on the conic which is tangent to $\overline{AC}$ at A, tangent to $\overline{BC}$ at B, and passes through a. Call this point c, then b is the reflexion of c in the line $\overline{Aa}$ with p as center.

*** It is of interest to note that the two imaginary intersections of these conics form with the four points A, B, C, a six-fold perspective triangles.**

Since the line $\overline{pc}$ ordinarily cuts the conic $x_1x_2 - x_3^2 = 0$ in two distinct points we find then on every line through p there are two sets of two points each which together with point a form a triangle four-fold perspective to ABC. There are two lines, however, through p which contain only one set of points. One is $\overline{BC}$ *; the other will be mentioned later.

Let us write the coördinates of the six points in terms of a parameter as follows:

$$\begin{array}{lll} A:\ (1,0,0), & B:\ (0,1,0), & C:\ (0,0,1), \\ a:\ (1,1,1), & b:\ (\mu^2,\mu,1), & c:\ (\mu^2,1,\mu). \end{array}$$

The coördinates of the four centers of perspection are $(\mu^2, 1, 1)$, $(\mu, 1, \mu)$, $(\mu, \mu, 1)$ and $(\mu, 1, 1)$. The condition that the first three centers lie on a line is

$$\begin{vmatrix} \mu^2 & 1 & 1 \\ \mu & 1 & \mu \\ \mu & \mu & 1 \end{vmatrix} = 0,$$

which reduces to $\mu(\mu - 1)^2(\mu + 2) = 0$. If μ equals 0 or 1, two points will coincide; hence *the condition that three centers of perspection lie on a line is* $\mu = -2$.

The condition that a conic can be put on the six vertices of the two triangles is

$$\begin{vmatrix} 1 & 1 & 1 \\ \mu^2 & 1 & \mu \\ \mu & 1 & \mu^2 \end{vmatrix} = 0,$$

which reduces to $\mu(\mu - 1)^2(\mu + 2) = 0$. The equation of the conic on the six points when $\mu = -2$ is $x_1x_2 - 2x_2x_3 + x_3x_1 = 0$; the equation of the line of centers is $x_1 - 2x_2 - 2x_3 = 0$; the odd center has coördinates $(-2, 1, 1)$. Hence the theorem *when* $\mu = -2$ *the six vertices of the two triangles lie on a conic, that three centers of perspection lie on a line and the fourth center of perspection is the pole of this line as to the conic.*

To construct the line of centers produce $\overline{Ba}$ to cut $\overline{AC}$ at D. On AC construct the fourth harmonic of C as to A and D; call this point q. Then $\overline{pq}$ is the required line.

Of the four intersections of the conics $x_1x_2 - 2x_2x_3 + x_3x_1 = 0$ and $x_1x_2 - x_3^2 = 0$ three are at A, B, a; the other is the point $(4, 1, -2)$. But this latter is the point of tangency of the tangent (other than $\overline{BC}$) from

* Obviously the set on BC coincides with B and C and the six points are not distinct.

p to $x_1x_2 - x_3^2 = 0$. This is the second exceptional line through p which contains only one set of two points which with the given point a makes a four-fold perspective triangle with ABC. From this we can readily deduce a construction for four-fold projective triangles whose vertices are on a conic. But a simpler method is available. We express the conic $x_1x_2 - 2x_2x_3 + x_3x_1 = 0$ in parametric form as

$$(5) \qquad x_1 = t(2 - t), \qquad x_2 = t, \qquad x_3 = 2 - t,$$

then the parameters of the points A, B, C are $t = \infty, 2, 0$ respectively while those of a, b, c are $t = 1, 4, -2$. Hence we have the theorem that *the six vertices of the two triangles on the conic are a cubic and its jacobian.* So the well known construction * on a conic for a cubic and its jacobian will furnish us four-fold perspective triangles whose vertices lie on a conic.

The Clebsch † six-point has the property that the 15 joins of the points meet by threes at ten points. The points can be arranged into two four-fold perspective triangles. The usual canonical form for the six-point is $1, 0, 0$; $1, 2\epsilon^4, 2\epsilon$; $1, 2\epsilon, 2\epsilon^4$; $1, 2, 2$; $1, 2\epsilon^2, 2\epsilon^3$ and $1, 2\epsilon^3, 2\epsilon^2$. The linear transformation which sends the first four of these points into the reference triangle and the unit point carries points 5 and 6 into $1 + \nu, 1, \nu$ and $1 + \nu, \nu, 1$ where $\nu = \frac{1}{2}(-1 + 5^{\frac{1}{2}})$. These points are on the line $-x_1 + x_2 + x_3 = 0$ which line is on p, the center of the reflexion and on $(1, 0, 1)$, the point where $\overline{Ba}$ cuts $\overline{AC}$. Points 5 and 6 are pairs in the reflexion with p as center and $\overline{Aa}$ as axis. To construct a Clebsch six-point with the apparatus available for four-fold perspective triangles means we must find the points where the conic cuts a definite line. This is a compass construction and the line can be constructed as indicated above. Since the conic cuts the line in two points there is another pair of points $5'$ and $6'$ which together with A, B, C, a will form a Clebsch six-point. The coördinates of this second set of points are $\nu, -1, 1 + \nu$ and $\nu, 1 + \nu, -1$ where $\nu = \frac{1}{2}(-1 + 5^{\frac{1}{2}})$.

For further study of four-fold perspective triangles let us choose a second canonical form. Let ABC be the reference triangle; in the previous form we let point a be the unit point and saw that it was on the conic $x_1x_2 - x_3^2 = 0$ so now we shall let point a run over this conic $x_1x_2 - x_3^2 = 0$ by giving it the coördinates $1, t^2, t$. Then point b will be on the conic $t^3x_1x_3 - x_2^2 = 0$ while point c lies on $x_1x_2 - x_3^2 = 0$. These two conics are reflected into each other by the reflexion with center $0, -t, 1$ and axis

* See for example Winger, *Projective Geometry*, p. 257.

† Clebsch, *Mathematische Annalen*, Vol. 4 (1871), p. 336.

$tx_3 - x_2 = 0$. The equations of the reflexion are $\rho x_1 = tx_1'$, $\rho x_2 = t^2 x_3'$, $\rho x_3 = x_2'$. The coördinates of the six points are now

$$A: (1, 0, 0), \qquad B: (0, 1, 0), \qquad C: (0, 0, 1),$$
$$a: (1, t^2, t), \qquad b: (t\mu^2, t^2\mu, 1), \qquad c: (\mu^2, 1, \mu).$$

The four centers of perspection are $t\mu^2, t, 1$; $\mu, t, t\mu$; $\mu, t^2\mu, 1$; and $\mu, t, 1$. The condition that the first three centers lie on a line is $t\mu + 2 = 0$. The equation of the line is then $t^2 x_1 - 2x_2 - 2tx_3 = 0$, whose envelope is $x_3^2 + 2x_1x_2 = 0$, a conic having double contact with $x_1x_2 - x_3^2 = 0$. The six vertices of the triangles now lie on the conic $tx_1x_2 - 2x_2x_3 + t^2x_3x_1 = 0$ and the odd center $-2, t^2, t$ has for its polar line as to this conic, the line on which lie the first three centers of perspection. In the general case, the odd center $\mu, t, 1$ and the center $t\mu^2, t, 1$ lie on the axis of reflexion $tx_3 - x_2 = 0$; the two remaining centers lie on a line through the center of reflexion $0, -t, 1$.

The axes of perspection have the following coördinates:

$$(6)\qquad \begin{array}{llll} A_1 \equiv A_aB_bC_c: & -t^2, & 1 + t\mu, & t(1 + t\mu), \\ A_2 \equiv A_bB_cC_a: & 1 + t\mu, & -\mu^2, & \mu(1 + t\mu), \\ A_3 \equiv A_cB_aC_b: & t(1 + t\mu), & \mu(1 + t\mu), & -t^2\mu^2, \\ A_4 \equiv A_aB_cC_b: & t, & \mu, & t\mu. \end{array}$$

The condition that the three axes A_1, A_2 and A_3 be on a point reduces to $(2t\mu + 1)(t\mu + 2) = 0$. If now $2t\mu + 1 = 0$, the three axes coincide and the vertices of the triangle abc are on a line. The second triangle has degenerated, but the centers of perspection are distinct and not on a line. Hence $t\mu + 2 = 0$ is the condition that the three axes meet in a point; it is likewise the condition that three centers of perspection be on a line and that the six vertices of the two triangles be on a conic. If, however, $2t\mu + 1 = 0$ the equation of the line on which the points a, b, and c lie is $2t^2x_1 - x_2 - tx_3 = 0$. The envelope of this line is $x_3^2 + 8x_1x_2 = 0$, another conic having double contact with $x_1x_2 - x_3^2 = 0$.

When $t\mu + 2 = 0$ the axes A_1, A_2 and A_3 meet at the point $2, -t^2, -t$; as t varies this point describes the conic $2x_3^2 + x_1x_2 = 0$. Likewise when $t\mu + 2 = 0$, the six sides of the two triangles touch a conic whose equation in line coördinates is $2tu_1u_2 - t^2u_2u_3 + 2u_3u_1 = 0$. The odd axis has coördinates $-t^2, 2, 2t$ and the three other axes meet on the point whose equation is $2u_1 - t^2u_2 - tu_3 = 0$. The pole of the odd axis as to the line conic is the point where the axes A_1, A_2 and A_3 meet. In general, the odd axis of perspection $tx_1 + \mu x_2 + t\mu x_3 = 0$ and the A_1 axis are on the center of re-

flexion $0, -t, 1$; while the axes A_2 and A_3 are on a point on the axis of reflexion $tx_3 - x_2 = 0$.

To summarize, when $t\mu + 2 = 0$, the six vertices of the two triangles lie on a conic; three centers of perspection are on the odd axis; three axes of perspection are on the odd center and this odd center and odd axis are pole and polar as to the conic. The three axes of perspection cut the odd axis in three points q_i which are the jacobian points of the three centers of perspection p_i on the line. The hessian points are the intersections of the line and conic. The three lines joining the points p_i to the odd center cut the conic in the six vertices of the two triangles; moreover the six tangents to the conic from the points q_i touch the conic at the same six vertices.

II. The Planar Group of Order 216.

On any two sides of the reference triangle in the plane choose a set of points such that the vertices of the triangle on those sides are the Hessian points of the set. Construct on both sides the cubicovariant points of each given set. Join now the six points on one side with the six points on the second side in all possible ways. These 36 lines will cut the third side of the triangle in six points having the same relation to it as the originally chosen sets have to the sides on which they are located. We thus have a set of 18 points and 36 lines such that 3 points are on each line and 6 lines on each point. The 36 lines fall naturally into four sets of nine each, such that each set together with the reference triangle form the flex triangles of a pencil of cubic curves.

If we designate by a_i the points $0, \omega^i, 1$; by b_i the points $\omega^i, 0, 1$; by c_i the points $\omega^i, 1, 0$; $(i = 0, 1, 2;\ \omega^3 = 1)$; and if we call the cubicovariant sets respectively by a_i', b_i' and c_i' one can easily verify that

$$(7)\quad \begin{array}{lll} a_i',\ b_i',\ c_i' & \text{are the flexes of} & x_1^3 + x_2^3 + x_3^3 + 6mx_1x_2x_3 = 0, \\ a_i',\ b_i,\ c_i & \text{are the flexes of} & -x_1^3 + x_2^3 + x_3^3 - 6mx_1x_2x_3 = 0, \\ a_i,\ b_i',\ c_i & \text{are the flexes of} & x_1^3 - x_2^3 + x_3^3 - 6mx_1x_2x_3 = 0, \\ a_i,\ b_i,\ c_i' & \text{are the flexes of} & x_1^3 + x_2^3 - x_3^3 - 6mx_1x_2x_3 = 0. \end{array}$$

The configuration consists of four Hesse configurations with a common triangle—the reference triangle— which is invariant under all the operations of the group. The group is the product of that subgroup G_{54} of a Hesse G_{216} which leaves one flex triangle invariant, and the G_4 which sends any one of the four Hesse configurations into each other. Hence we have an imprimitive G_{216}.

Now these 36 lines intersect in 360 points outside of the points on the

sides of the reference triangle. These 360 points divide into three sets containing 36, 108 and 216 respectively. For each line is conjugate under the G_{216} and hence is unaltered by a G_6. If two of the 36 lines should be unaltered by the same G_3, then the G_2 which interchanges them will leave their join unaltered by a G_6. Hence a set of 36 conjugate points. This will occur if the two lines belong to a set of flex lines of a cubic. This is evident also from the duality existing. For the 18 lines $x_1 \pm \omega^i x_2 = 0$, $x_2 \pm \omega^i x_3 = 0$, $x_3 \pm \omega^i x_1 = 0$ $(i = 0, 1, 2)$ which go by six through the vertices of the reference triangle meet in the 36 points $1, \pm\omega^i, \pm\omega^j$; $(i = j = 0, 1, 2)$. These are the coördinates of the set of 36 conjugate points mentioned above. These 36 points are on 360 lines, outside of those passing through the vertices of the reference triangle. These 360 lines fall into three sets containing 36, 108 and 216 respectively. The set of 36 lines are the flex lines of the four cubics; the 36 points are the flexes.

Furthermore, if two of the original 36 lines are not invariant under the same G_3, they may be interchanged by a G_2 or not. In the former case we shall get a set of 108 conjugate points, in the latter case a set of 216 points conjugate under the group. These will be discussed later, the dual sets of 108 and 216 lines will also appear.

Now the 36 points $1, \pm\omega^i, \pm\omega^j$ $(i = j = 0, 1, 2)$ fall into twelve triangles. Let

$$
\begin{array}{lllll}
A_{11} & \text{be} & 1, 1, 1; & 1, \omega, \omega^2; & 1, \omega^2, \omega; \\
A_{12} & \text{be} & 1, 1, \omega^2; & 1, \omega^2, 1; & 1, \omega, \omega; \\
A_{13} & \text{be} & 1, 1, \omega; & 1, \omega, 1; & 1, \omega^2, \omega^2; \\
A_{21} & \text{be} & 1, -1, 1; & 1, -\omega, \omega^2; & 1, -\omega^2, \omega; \\
A_{22} & \text{be} & 1, -1, \omega^2; & 1, -\omega^2, 1; & 1, -\omega, \omega; \\
A_{23} & \text{be} & 1, -1, \omega; & 1, -\omega, 1; & 1, -\omega^2, \omega^2; \\
A_{31} & \text{be} & 1, 1, -1; & 1, \omega, -\omega^2; & 1, \omega^2, -\omega; \\
A_{32} & \text{be} & 1, 1, -\omega^2; & 1, \omega^2, -1; & 1, \omega, -\omega; \\
A_{33} & \text{be} & 1, 1, -\omega; & 1, \omega, -1; & 1, \omega^2, -\omega^2; \\
A_{41} & \text{be} & -1, 1, 1; & -1, \omega, \omega^2; & -1, \omega^2, \omega; \\
A_{42} & \text{be} & -1, 1, \omega^2; & -1, \omega^2, 1; & -1, \omega, \omega; \\
A_{43} & \text{be} & -1, 1, \omega; & -1, \omega, 1; & -1, \omega^2, \omega^2.
\end{array} \tag{8}
$$

We shall arrange these twelve triangles in four rows of three each as follows:

$$
\begin{array}{ccc}
A_{11} & A_{12} & A_{13} \\
A_{21} & A_{22} & A_{23} \\
A_{31} & A_{32} & A_{33} \\
A_{41} & A_{42} & A_{43}.
\end{array}
$$

Now this set of twelve triangles has the following remarkable property—*any two triangles in the same row are six-fold perspective; any two triangles in the same column are four-fold perspective with three centers of perspection on a line;* finally *any two triangles in different rows and columns are four-fold perspective.* Since the triangles in any row together with the reference triangle are the flex triangles of a cubic we know they are six-fold perspective by pairs, with the six centers of perspection being the other six points.

Any two triangles in the same column are four-fold perspective with three centers of perspection on a line. For A_{11} and A_{21} the centers of perspection are $1, 0, \omega^i$; $(i = 0, 1, 2)$ and $0, 1, 0$; the three flexes on one side of the reference triangle and the opposite vertex. Since the G_{216} is doubly transitive on the triangles in a column, this proves the theorem for any two triangles in any column. The triangles can be paired thus in 18 ways. In each case the three centers of perspection on a line, lie on the sides of the reference triangle. The three axes of perspection on a point for A_{11} and A_{21} are $x_1 + \omega^i x_3 = 0$ $(i = 0, 1, 2)$ on the point $0, 1, 0$; the odd axis is the opposite side of the reference triangle $x_2 = 0$. Moreover $A_{11}A_{21}$, $A_{12}A_{22}$, and $A_{13}A_{33}$ have the same centers and same axes of perspection. If two triangles are four-fold perspective with three centers on a line, the six vertices lie on a conic. These conics are

$$(9)\qquad \begin{aligned} A_{1i}A_{2i}&: & x_2^2 - \omega^{i-1}x_1x_3 &= 0 \\ A_{1i}A_{3i}&: & x_3^2 - \omega^{i-1}x_1x_2 &= 0 \\ A_{1i}A_{4i}&: & x_1^2 - \omega^{i-1}x_2x_3 &= 0 \\ A_{2i}A_{3i}&: & x_1^2 + \omega^{i-1}x_2x_3 &= 0 \\ A_{2i}A_{4i}&: & x_3^2 + \omega^{i-1}x_1x_2 &= 0 \\ A_{3i}A_{4i}&: & x_2^2 + \omega^{i-1}x_1x_3 &= 0 \end{aligned} \qquad (i = 1, 2, 3).$$

These 18 conics belong to three pencils of double contact conics, the parameters of the six conics in each pencil being the sixth roots of unity. The chord of contact of each pencil is one side of the reference triangle. Each conic is unaltered by a G_{12} of the group.

Two triangles in a different row and column are four-fold perspective. Thus the four centers of perspection for the triangles A_{13} and A_{22} are $0, 1, 0$; $1, 1 - \omega^2, \omega^2$; $1, \omega - 1, \omega$; and $1, \omega^2 - \omega, 1$. Two such triangles can be joined in 72 ordered pairs or 36 non-ordered pairs. The G_{18} which leaves A_{13} invariant is generated by

$$(10)\qquad \begin{aligned} \rho x_1' &= x_1, & &= x_3, & &= x_1 \\ \rho x_2' &= x_2, & &= x_1, & &= x_3 \\ \rho x_3' &= \omega^2 x_3, & &= x_2, & &= x_2. \end{aligned}$$

This sends triangle A_{22} into itself, or A_{32} or A_{42}. The G_{216} sends A_{13} into each of the twelve triangles, and carries its partner A_{22} along into three, consequently we get 36 non-ordered pairs. Hence any pair of triangles can be sent by the group into any other pair in a certain order. Thus all 36 pairs are four-fold perspective. The G_6 leaving both A_{13} and A_{22} unaltered is generated by

$$(11) \qquad \begin{aligned} \rho x_1' &= x_1, & &= x_3 \\ \rho x_2' &= \omega x_2, & &= x_2 \\ \rho x_3' &= \omega x_3, & &= x_1. \end{aligned}$$

This G_6 leaves the center of perspection $(0, 1, 0)$, a vertex of the reference triangle, unaltered; it permutes the other three centers in all possible ways. We have thus $36 \times 3 = 108$ points, conjugate under the group. These points can be identified as that set of 108 conjugate points mentioned earlier in this section. Similarly the 108 axes of perspection of these 36 non-ordered pairs of triangles, outside of the sides of the reference triangle, are the set of 108 conjugate lines previously mentioned.

Under the G_{216} points fall into sets of 216 conjugates points. But this particular set of 216 points which appeared earlier is worth noticing. They lie by 12's on 36 lines. Thus on $-x_1 + x_2 + x_3 = 0$ we have the pair $\omega^2 - \omega, 1, 2\omega^2$; $\omega^2 - \omega, 2\omega^2, 1$. The remaining five pairs of points are given by that G_6 which sends the line into itself.

$$(12) \qquad \begin{aligned} \rho x_1 &= x_1', & &= x_2', & &= x_3', & &= x_1', & &= x_3', & &= x_2' \\ \rho x_2 &= x_2', & &= -x_3', & &= x_1', & &= x_3', & &= -x_2', & &= x_1' \\ \rho x_3 &= x_3', & &= x_1', & &= -x_2', & &= x_2', & &= x_1', & &= -x_3'. \end{aligned}$$

Let us consider the six-point $1, 0, 0$; $0, 1, 0$; $0, 0, 1$; $1, 1, 1$; $\omega^2 - \omega, 1, 2\omega^2$; $\omega^2 - \omega, 2\omega^2, 1$. Coble * has given, in a paper on Point Sets and Cremona Groups, a method for calculating certain irrational invariants of a six-point which he has called $\bar{a}, \bar{b}, \cdots \bar{f}$ where $\sum^{6} \bar{a} = 0$. If two of these invariants become equal, certain lines on the six points meet by threes. For our special six-point $\bar{a} = \bar{c} = \bar{d} = \bar{e}$. This means geometrically that points 5 and 6 are conjugate points in the reflexion set up with center at meet of $\overline{23}$ and $\overline{56}$ and with axis as $\overline{14}$; also that the pair of points 1 and 4 are apolar to the conic $\overline{23}, \overline{56}$. Consequently the six-point is self associated in the order (2536). Further the double ratio of the four lines 1-2356 or of 4-2356 has the value 4ω; similarly the double ratio of the lines 5-1234, or of 6-1234, or of 2-1456,

* A. B. Coble, *Transactions of the American Mathematical Society*, Vol. 16 (1915). p. 155.

or of 3-1456 is 4ω. Certain of the lines on the six-point meet by threes in six points whose coördinates are $\omega^2-\omega, 1, 1$; $1-\omega^2, 2, 2$; $1, 1, 0$; $0, 2\omega^2, 1$; $0, 1, 2\omega^2$; $1, 0, 1$. It is interesting to note this new six-point is of the same type as the first one, since four of the six irrational invariants of it are equal.

Now on the line $-x_1+x_2+x_3=0$ are six pairs of points; each pair together with the reference triangle and one other point from $1, 1, 1$; $1, -1, 1$; or $1, 1, -1$ forming a six-point of the type above. Hence the set of 216 points fall into pairs such that each pair with the reference three-point and a definite flex triangle vertex forms a six-point of the above nature. We have thus 216 such six-points, with each of the 36 flex triangle vertices used six times.

A PREPARED SYSTEM FOR TWO QUADRATICS IN SIX VARIABLES.

By J. Williamson.

Introduction. In a previous paper,* a prepared system was determined, in terms of which every concomitant of two quadratics in n variables could be expressed, if the concomitants were multiplied by suitable invariant factors. In this paper a prepared system, for the case $n = 6$, is determined, in terms of which every concomitant can be expressed, without being multiplied by an invariant factor. It is found that 52 new factors must be added to the $2^6 - 1 = 63$ factors already determined, making a total of 115.

The notation of the previous paper is used throughout except that, for convenience in printing, dashes are used instead of dots to denote determinantal permutations; i. e. the series $(abc)d_x - (abd)c_x - (adc)b_x$ is denoted by one of the three expressions $(ab'c')d_x'$, $(ab''c'')d_x''$, $(ab'''c''')d_x'''$. In addition, for the six sets of cogredient point variables, that are necessary for this discussion, x, y, z, t, w, k are now used, while Q, P, p, and u are written for the compound coördinates π_2, π_3, π_4 and π_5 respectively.†

The first section gives a list of the results while the second is devoted to their determination.

1. *The Prepared System.* The prepared system consists of 115 factors. Of these 63 are simple factors;

$$\left.\begin{array}{l}6\ x\text{-factors of type } i_x,\\ 6\ u\text{-factors of type } (jkmnt),\end{array}\right\}\ \text{duals;}$$

$$\left.\begin{array}{l}15\ Q\text{-factors of type } (ij),\\ 15\ p\text{-factors of type } (kmnt),\end{array}\right\}\ \text{duals.}$$

20 P-factors of type (ijk), 1 factor (123456).

47 are linear in two sets of variables;

3 ux-factors of type $i_x'(j'kmnt)$,

3 pQ-factors of type $(i'j')(k'mnt)$,

$$\left.\begin{array}{l}8\ Qu\text{-factors of type } (ij')(imntk')\\ 8\ px\text{-factors of type } (mntk')j_x',\end{array}\right\}\ \text{duals.}$$

* J. Williamson, "A Special Prepared System for Two Quadratics in n Variables," *American Journal of Mathematics*, Vol. 52 (April, 1930), pp. 399-412.

† *Loc. cit.*, §§ 1 and 2.

6 Pu-factors of type $(ijk')(ijt'nm)$ } duals;
6 Px-factors of type $(mnt')k_x'$,
12 pQ-factors of type $(ij')(k'imn)$
1 pQ-factor $(123'5'')(4'6'')$.
1 is quadratic in the variable P: $(123')(4'56)$.
1 is quadratic in x and linear in p: $(123'5'')4_x'6_x''$. } duals.
1 is quadratic in u and linear in Q: $(4''6')(123''56)(12345')$
2 are linear in the three variables P, u and x;
$(1''6'3)(2''6543)5_x'$, } duals.
$(5''42')(12346'')1_x'$,

These factors illustrate very clearly how the principle of duality * applies to the non-simple bracket factors. Corresponding to every non-simple factor is a dual factor formed by taking the duals of the component factors and permuting the same symbols. For example, $1_x'(2'34)$ yields the dual factor $(2'3456)(1'56)$. A factor may of course be self dual, as is the case with $(123')(4'56)$.

A complete list of these 115 factors is given below. In this list I denotes a product of invariant factors formed from a_ρ, (AR_4), (A_3R_3), (A_4R), r_α, where A, R, α and ρ are written for A_2, R_2, A_5 and R_5 respectively. If in a factor 12 is convolved, I includes a_ρ, if 23, (AR_4) etc. and in any particular case the value of I may be written down immediately. If two factors are similar,† only one has been defined, since the other may be obtained by replacing a, A, A_3, A_4, α by r, R, R_3, R_4, ρ respectively.

List of factors.

$1_x = a_x$, 6_x, $2_x = (A\rho x) = a_\rho' b_x'$, 5_x,
$3_x = (A_3R_4x) = (a'b'R_4)c_x'$, 4_x;
$(12) = a_\rho(AQ)$, (65), $(13) = (aA_3R_4Q) = (a'b'R_4)(ac'Q)$, (64),
$(14) = (aA_4R_3Q) = (a'b'c'R_3)(ad'Q)$, (63),
$(15) = (aR\alpha Q) = r_\alpha'(as'Q)$, (62), $(16) = (arQ)$,
$(23) = (AR_4)(A_3\rho Q) = (AR_4)a_\rho'(b'c'Q)$, (54),
$(34) = (A_3R_3)(A_4R_4Q) = (A_3R_3)(a'b'R_4)(c'd'Q)$,
$(24) = (A\rho A_4R_3Q) = (A_4r''s'')(b't''Q)a_\rho'$, (53),
$(25) = (A\rho R\alpha Q) = a_\rho'' r_\alpha'(b''s'Q)$;
$(123) = I(A_3P)$, (654),
$(124) = I(AA_4R_3P) = I(A_4r's')(At'P)$, (653),
$(125) = I(AR\alpha P) = Ir_\alpha'(As'P)$, (652), $(126) = I(ArP)$, (651),

* *Loc. cit.*, § 5.
† *Loc. cit.*, § 6.

$(134) = I(aA_4R_4P) = I(a'b'R_4)(ac'd'P)$, (643),
$(135) = (aA_3R_4R\alpha P) = (a''b''R_4)r_a'(ac''s'P)$, (642),
$(136) = (aA_3R_4rP) = (a'b'R_4)(ac'rP)$, (641),
$(154) = I(aR_3\alpha P) = Ir_a'(as't'P)$, (623),
$(234) = I(A_4\rho P) = Ia_\rho'(b'c'd'P)$, (543),
$(254) = I(A\rho R_3\alpha P) = Ia_\rho'(b'R_3\alpha P)$, (523);
$(6543) = I(R_4p)$, (1234), $(6542) = I(R_3A\rho p) = I(R_3b'p)a_\rho'$, (1235),
$(6523) = I(RA_3\rho p) = Ia_\rho'(Rb'c'p)$, (1254),
$(6234) = I(rA_4\rho p) = Ia_\rho'(rb'c'd'p)$, (1543),
$(2345) = I(\alpha\rho p) = Ia_\rho'(b'c'd'e'p)$, $(1654) = I(aR_3p)$, (6123),
$(1365) = I(aA_3R_4Rp) = I(a'b'R_4)(ac'Rp)$, (6412),
$(1346) = I(aA_4R_4rp) = I(r's'A_4)(ak't'p)$, $(1265) = I(ARp)$;
$(12345) = I(\alpha u)$, (65432), $(12346) = I(A_4ru)$, (65431),
$(12365) = I(A_3Ru)$, (65412);
$(123456) = a_\rho(AR_4)(A_3R_3)(A_4R)(\alpha r)$;
$(12, 6543) = 1_x'(2'6543) = -6_x'(125'4'3') = I(AR_4ux) = Ia_x'(b'R_4u)$,
$(65, 1234) = 6_x'(5'1234) = -1_x'(652'3'4')$,
$(123, 654) = 1_x'(2'3'654) = 6_x'(1235'4') = I(A_3R_3ux) = Ia_x'(b'c'R_3u)$;
$(123, 1654) = (12')(3'1654) = (16')(1235'4') = I(A_3aR_3Qu) = I(a'b'Q)(c'aR_3u)$,
$(654, 6123) = (65')(4'6123) = (61')(6542'3')$,
$(165, 1234) = (16')(5'1234) = (13')(6514'2') = I(aRA_4Qu) = I(ar'Q)(s'A_4u)$,
$(612, 6543) = (61')(2'6543) = (64')(1263'5')$,
$(265, 1234) = (26')(5'1234) = (3'2)(65124') = I(A\rho RA_4Qu) = I(A\rho r'Q)(s'A_4u)$,
$(512, 6543) = (51')(2'6543) = (4'5)(12653')$,
$(123, 6543) = (1'3)(2'6543) = (6'3)(215'4'3) = I(A_3R_4Qu) = I(a'c'Q)(b'R_4u)$,
$(654, 1234) = (6'4)(5'1234) = (1'4)(562'3'4)$,
$(1265, 6543) = (1'65)(2'6543) = (4'65)(21653')$,
$= I(ARR_4Pu) = I(a'RP)(b'R_4u)$.
$(6512, 1234) = (6'12)(5'1234) = (3'12)(56124')$.
$(1236, 6543) = (1'36)(2'6543) = (5'36)(1264'3)$,
$= I(A_3rR_4Pu) = I(a'c'rP)(b'R_4u)$,
$(6541, 1234) = (6'41)(5'1234) = (2'41)(6513'4)$,
$(1234, 6543) = (1'34)(2'6543) = (5'34)(126'43)$,
$= I(A_4R_4Pu) = I(a'c'd'P)(b'R_4u)$,
$(1236, 6541) = (12'6)(3'6541) = (15'6)(3624'1)$,
$= I(A_3rR_3aPu) = I(a'b'rP)(c'R_3au)$;
$(123, 654) = (1'2')(3'654) = (6'5')(1234')$,
$= I(A_3R_3Qp) = I(a'b'Q)(c'R_3p)$,
$(126, 543) = (1'6)(2'543) = (5'6)(214'3')$,
$= I(ArR_4\alpha Qp) = I(a'rQ)(b'R_4\alpha p)$,
$(651, 234) = (6'1)(5'234) = (2'1)(563'4')$,

$$(123, 154) = (12')(3'154) = (15')(1234'),$$
$$= I(A_3 a R_3 \alpha Q p) = I(a'b'Q)(c'aR_3\alpha p) = I(a'b'Q)(c'as''t''p)r_a'',$$
$$(654, 623) = (65')(4'623) = (62')(6543'),$$
$$(123, 165) = (13')(2'165) = (15')(1236') = I(A_3 a R Q p),$$
$$= I(a'c'Q)(b'aRp) = -I(ar'Q)(A_3 s'p),$$
$$(654, 612) = (64')(5'612) = (62')(6541'),$$
$$(134, 165) = (13')(4'165) = (16')(4135'),$$
$$= I(aA_4R_4aRQp) = I(ar'Q)(aA_4R_4s'p),$$
$$(643, 612) = (64')(3'612) = (61')(3642'),$$
$$(234, 265) = (23')(4'265) = (26')(2345') = I(A_{4\rho}A_\rho RQp),$$
$$= I(b'c'Q)(d'A_\rho Rp)a_\rho' = I(b'c'Q)(d'Ab''p)a_\rho' a_\rho'',$$
$$(543, 512) = (54')(3'512) = (51')(5432'),$$
$$(123, 543) = (1'3)(2'543) = (5'3)(214'3),$$
$$= I(A_3R_4\alpha Qp) = I(a'c'Q)(b'R_4\alpha p),$$
$$(654, 234) = (6'4)(5'234) = (2'4)(563'4),$$
$$(123, 365) = (1'3)(2'365) = (6'3)(1235') = I(A_3A_3R_4RQp),$$
$$= I(a'c'Q)(b'A_4R_3Rp) = I(r'A_3R_4Q)(A_3s'p),$$
$$(654, 412) = (64')(5'412) = (1'4)(6542'),$$
$$(12, 34, 65) \doteq (123'6'')(4'5'') = (1'436'')(2'5'') = (1'53''6)(2'4''),$$
$$= I(AA_4R_4RQp) = I(Ac'r''p)(d's''Q)(a'b'R_4);$$
$$(12, 543) = 1_x'(2'543) = 4_x'(125'3') = I(AR_4\alpha px) = Ia_x'(b'R_4px),$$
$$(65, 234) = 6_x'(5'234) = 3_x'(652'4'),$$
$$(12, 346) = 1_x'(2'346) = 4_x'(123'6) = I(AA_4R_4rpx) = Ia_x'(b'A_4R_4rp),$$
$$(65, 431) = 6_x'(5'431) = 3_x'(654'1),$$
$$(12, 654) = 1_x'(2'654) = (126'4')5_x' = I(AR_3px) = Ia_x'(b'R_3p),$$
$$(65, 123) = 6_x'(5'123) = 2_x'(651'3'),$$
$$(23, 654) = 2_x'(3'654) = 5_x'(236'4') = I(A_{3\rho}R_3px) = Ia_\rho' b_x'(c'R_3p),$$
$$(54, 123) = 5_x'(4'123) = 2_x'(541'3');$$
$$(12, 34) = 1_x'(2'34) = 4_x'(123') = I(AA_4R_4Px) = Ia_x'(b'A_4R_4P),$$
$$(65, 43) = 6_x'(5'43) = 3_x'(654'),$$
$$(12, 65) = 1_x'(2'65) = 5_x'(126') = I(ARPx) = Ia_x'(b'RP),$$
$$(23, 54) = 2_x'(3'54) = 4_x'(235') = I(A_{3\rho}R_3\alpha Px) = Ia_\rho' b_x'(c'R_3\alpha P),$$
$$(12, 54) = 1_x'(2'54) = 4_x'(125') = I(AR_3\alpha Px) = Ia_x'(b'R_3\alpha P),$$
$$(65, 23) = 6_x'(5'23) = 3_x'(652');$$
$$(12, 34, 65)' = (123'6'')4_x'5_x'' = (1'436'')2_x'5_x'' = (1'4''56)2_x'3_x'',$$
$$= I(AA_4R_4Rpxx) = I(Ac'r''p)(a'b'R_4)d_x's_x'';$$
$$(1265, 6543, 1234) = (1'6'')(2'6543)(5''1234),$$
$$= I(AR_4RA_4Quu) = I(a'r''Q)(b'R_4u)(s''A_4u);$$
$$(123, 6543, 65) = (1'6''3')(2'6543)5_x'',$$
$$= I(A_3R_4RPux) = I(a'r''c'P)(b'R_4u)s_x'',$$
$$(654, 1234, 12) = (6'1''4')(5'1234)2_x'';$$
$$(12, 34, 65)'' = (123')(4'65) = (1'43)(2'65) = (126')(435'),$$
$$I(AA_4R_4RPP) = I(Ac'P)(d'RP)(a'b'R_4).$$

In finding the values of I in the above list symbols separated by a comma are not to be counted as convolved; thus in $(12, 34)$ the value of I is $a_\rho(A_3R_3)$ and in $(12, 34, 65)$ is $a_\rho(A_3R_3)r_a$. Throughout we have written

$$A = ab, \quad A_3 = abc, \quad A_4 = abcd, \quad \alpha = abcde, \quad R = rs, \quad R = rst, \quad R_4 = rstk.$$

In the definitions of the bracket factors, it is sometimes necessary to use previous definitions; for example,

$$(12, 34) = Ia_x'(b'A_4R_4P) = Ia_x'(a''b''R_4)(b'c''d''P),$$

by the definition of (134).

2. *Determination of the Prepared System.* Since we are now considering two quadratics in n variables for the case $n = 6$, there are seven invariants and six quadratic covariants i_x^2, $(i = 1, 2, 3, 4, 5, 6)$.* By theorem I every concomitant,† multiplied by a suitable invariant factor, can be expressed in terms of the symbolic factors,

$$i_x, \ (ij), \ (ijk), \ (ijkm), \ (ijkmn), \ (123456) \qquad (i, j, k, m, n = 1, 2, 3, 4, 5, 6).$$

We must now determine if ever in forming these bracket factors, we have disturbed any of the invariant factors, which appear when $12, 23, 34, 45, 56$ are convolved together. Originally we have six sets of cogredient point variables x, y, z, t, w, k, which are convolved as $\Delta = (xyztwk)$, $u = xyztw$, $p = xyzt$, $P = xyz$, $Q = xy$. Since the only factor involving all six variables is (123456), and in this $12, 23, 34, 45, 56$ are convolved, it follows that no invariant factor is disturbed in forming Δ. Let us consider the formation of the factors involving the variable u first of all. For simplicity we call a factor containing m symbols an m-factor. If one of the variables x, y, z, t, w convolved to form u appear in a four-factor, we may take four of these variables as appearing in that factor. For ‡

$$(ijkr \mid xyzt')(mn \mid w'y) = (ijkr \mid x'y'wt)(mn \mid z'y) + (ijkrm' \mid u)n_y',$$

and on the right w and t are both convolved in the same factor. Proceeding in this way we see that we lose nothing by assuming that four of the variables occur in the four-factor. We have then to consider the cases in which the fifth variable occurs in a two-factor, a three-factor or a four-factor. The

* *Loc. cit.*, p. 404.

† *Loc. cit.*, § 3.

‡ Both here and later the symbols $i, j, k, t, m, n, a, b, c, d, e, f$ are used to denote any of the symbols 1, 2, 3, 4, 5, 6.

case with the fifth variable in a one-factor obviously does not require to be considered. If the four-factor is $(ijkr \mid xyzt)$ or more shortly $(ijkr)$, the two-factor cannot contain any of the symbols i, j, k, r; the three-factor can contain at most one and the four-factor at most two of the symbols i, j, k, r, for otherwise in the formation of u no invariant factors would be disturbed. Since there are only six possible values for i, j, k, r, we must consider

$$(ijkrn')m', \quad (ijkrn')(im'), \quad (ijkrn')(ijm'),$$

where i, j, k, m, n are all distinct and we have not written in the variables.

If none of the variables, convolved to form u, occur in a four-factor, one may occur in a three-factor. In this case, as before, we may assume that three of the variables forming u occur in this factor and we have to consider the cases: (a) three variables in one three-factor, two in another; (b) three variables in one three-factor, one in each of two three-factors; (c) three variables in one three-factor, one in another three-factor, one in a two-factor; (d) three variables in one three-factor, one in each of two two-factors. In case (a) the same symbol cannot appear in both three-factors and accordingly we have the sole possibility $(ijkr'm')n'$. In case (b) we have

$$(ijkr'a'')(m'n')(b''c''),$$

where no two of i, j, k are the same as two of r, m, n or of a, b, c nor two of a, b, c are the same as two of r, m, n. For, if $rmn = ijk$, (b) becomes $(ijkna'')(ij)(b''c'')$ and here ijn are still convolved. The rest follows from the fact that we might have started with the factor (rmn) or (abc) in place of (ijk). In case (c) we have $(ijkr'a'')(m'n')b''$, where $ijkrmn$ must involve at least five distinct symbols. But neither of a, b can be the same as one of i, j, k or the same as one of r, m, n and hence this type is impossible. In case (d) we have $(ijkr'a'')m'b''$, where r, m and a, b contain no symbols in common with i, j, k. If $ab = rm$, this type is obviously reducible * and so we must consider the case

$$(ijkr'r'')m'n'' = (ijknr)m'r' + (i'j'mnr)k'r$$

and each term on the right reduces to simpler bracket factors. The further case, in which only two-factors can occur is easily seen to be impossible.

If the variable w does not occur but the variable t does, that is if the

* We use the phrase "is reducible" to denote that the factor under consideration can be expressed in terms of simpler types or of types that have already been considered. The sign $\equiv$ is used for "equal to, apart from reducible terms."

coördinate p appears but not the coödinate u, and one of the three variables convolved to form p occur in a three-factor, three of them may be considered to occur in that factor. Hence we have the types, $(ijkr')(m'n')$ from two three-factors, $(ijkr')(im')$ from two three-factors, $(ijkr')m'$ from one three-factor and one two-factor. But, if no three-factor occur, we have the type $(ijk'm'')r'n''$ from three two-factors.

If no variables w or t occur, but z occurs, we have the single type $(ijk')r'$. Accordingly we have to consider the following types.

A. $(ijkrm')n'$, B. $(ijkrm')(in')$, C. $(ijkrm')(ijn')$,
D. $(ijkr'm')n'$, E. $(ijkr'n'')(im')(jm'')$, F. $(ijkr')(m'n')$,
G. $(ijkr')(im')$, H. $(ijkr')m'$, I. $(ijk'm'')r'n''$,
J. $(ijk')m'$.

We now consider these types in detail.

Type A. In A m, n must be successive integers and so we have the possibilities,

$$1'(2'3456),\quad 2'(3'1456)=6'(1234'5')+1(32456),$$
$$3'(4'1256)=1'(2'3456)+6'(12345'),$$
$$4'(5'1236)=3'(4561'2')+6(51234),\quad 5'(6'1234).$$

We are accordingly left with only two of this type, $1'(2'3456)$ and $5'(6'1234)$, if we include type D. For example the term $1(32456)$ on the right of $2'(3'1456)$ can be neglected, since (32456) has $32, 45, 56$ all convolved.

Type B. In type B m, n must be successive integers and by letting $i=1, 2, 3$ in turn we have the possibilities;

$$(12')(14563'),\quad (13')(12564')\equiv(15')(12346'),$$
$$(14')(12365')\equiv(12')(14563'),\quad (15')(12346'),$$
$$(23')(21564')\equiv(25')(12346'),\quad (24')(21365')\equiv(26)(21345),$$
$$(25')(21346'),\quad (32')(34561'),\quad (34')(32165')\equiv(32')(34561'),$$
$$(35')(32146')\equiv(32')(34561'),$$

together with similar types. These types reduce as indicated above to the eight,

$$(12')(14563'),\quad (65')(63214'),\quad (15')(12346'),\quad (62')(65431'),$$
$$(25')(21346'),\quad (52')(56431'),\quad (31')(34562'),\quad (46')(43215').$$

Type C. In type C, m, n must be successive integers and so must k, r.

Accordingly m, n and k, r may have the following values: $m, n = 1, 2$; $k, r = 3, 4$ or $4, 5$ or $5, 6$: $m, n = 2, 3$; $k, r = 4, 5$ or $5, 6$: $m, n = 3, 4$; $k, r = 5, 6$. From these values we obtain six types,

$$(1'56)(56342'), \quad (1'63)(63452'), \quad (1'34)(34562'),$$
$$(2'16)(16453'), \quad (2'14)(14563'), \quad (4'12)(12563').$$

Type D. Since r, m, n and i, j, k must be successive integers, there is only one factor of this type, $(1234'5')6'$.

Type E. Since

$$(ijkr'n'')(im')(jm'') \equiv (ijkrm)(in'')(jm'') + (ij'mrn'')(ik')(jm''),$$
$$\equiv (ij'mrn'')(ik')(jm''),$$

for the other term contains the simple factor $(ijkrm)$, we may interchange the rôles of ijk and imr, and similarly the rôles of ijk and jmn. If we consider the factors ijk, imr, jmn in turn as the foundation for the u factor, we see that ijk, imr, jmn must all be sets of three successive integers. They must be chosen from 123, 234, 345, 456 and any three of these sets include two with two symbols the same. Accordingly a factor of type E would simplify.

Type F. In this type at least two of i, j, k and at least two of r, m, n must be successive integers and so we have the possible cases,

$$(1234')(5'6'), \quad (1243')(5'6') \equiv (1245')(36'),$$
$$(1253')(4'6') \equiv (1253')(4'6) \equiv (1'345)(62'),$$
$$(1263')(4'5') \equiv (1'345)(62'),$$
$$(1342')(5'6') \equiv (1345')(26') \equiv (1563')(4'2) \equiv (1562')(3'4'),$$
$$(1452')(3'6') \equiv (1452')(3'6) \equiv (1234')(5'6),$$
$$\equiv (1234')(5'6') - (1236)(45),$$
$$(1562')(3'4') \equiv (2345')(16').$$

These factors reduce as indicated above to the three,

$$(1234')(5'6'), \quad (3452')(61'), \quad (4325')(16').$$

Type G. In this type m, n must be successive integers and so must j, k. We let $i = 1, 2, 3$ in turn and so get the six factors,

$$(12')(13'45), \quad (12')(13'56), \quad (13')(14'56),$$
$$(23')(24'56), \quad (31')(32'45), \quad (31')(32'56),$$

and six similar factors making twelve in all.

Type H. In this type r, m must be successive integers and so must at least two of i, j, k. We accordingly have

$$\begin{array}{llll} 1'(2'345), & 1'(2'346), & 1'(2'356) \equiv 6'(1235'), & 1'(2'456), \\ 2'(3'145) \equiv (1235')4', & & 2'(3'156) \equiv (1236')5', & 2'(3'456), \\ 3'(4'126) \equiv 1'(2'346), & & 3'(4'125) \equiv 1'(2'345), & \\ 3'(4'156) \equiv 6'(1345'), & & 3'(4'256) \equiv 6'(2345'), & \end{array}$$

and similar factors. But these reduce as indicated above to the four factors,

$$1'(2'345), \quad 1'(2'346), \quad 1'(2'456), \quad 2'(3'456),$$

and four similar factors.

Type I. In this type i, j; k, r; m, n must all be distinct and must all be pairs of successive integers. Accordingly there is only the one possibility $(123'5'')4'6''$.

Type J. In this type i, j and k, m must both be pairs of successive integers and so we have the types.

$$1'(2'34), \quad 1'(2'45), \quad 1'(2'56), \quad 2'(3'45), \quad 2'(3'56), \quad 3'(4'56).$$

Further u-factors. Since types F, G, H, I and J only arise when no variable u is present, we need only consider types A, B, C, and D in the formation of new u-factors. Let us first consider $C = (ijkrm')(ijn' \mid Y)$. If one of the variables of Y is to be convolved to form u and none of the components of u occur in a four-factor, then all three variables in Y may be taken as forming part of u. Thus we have the possibilities, $(ijkrm')(ijn'ab)$, $(ijkrm')(ijn'a''b)$, $(ijkrm')(ijn'a''b'')c''$, $(ijkrm')(ijn'a''b''')$.* In the first case neither of $a, b = i, j, m$ or n and so ab must be kr. By the fundamental identities † this reduces. Similarly the second and the fourth obviously reduce. In the third case, none of a, b, c is i or j; hence two of them are m, n or k, r and if $ab = mn$, $c = k$. We have then the type

$$(ijkrm')(ijn'm''r'')n'',$$

which is reducible. Since k, r and n, m are interchangeable in C, this type reduces in every case. But, if one of the variables of u occur in a four-factor, we have to consider the possibility $(ijkrm')(i'j')(n'abcd)$. This is obviously reducible, if $abcd$ includes ij. Let now $abcd$ involve i but not j, then since

* The symbols ′ ″ ‴ after a letter mean that the letter so marked belongs to a convolution of letters even though the other members of the convolution are omitted.

† *Loc. cit.*, p. 408.

k, r and m, n are interchangeable in C, $(ijkrm')(ij')(n'imnr)$ is typical. But this last is equivalent to $(ijkrm')(in')(jimnr)$, which is a product of factor types already considered. We are left to consider $(ijkrm')(i'j')(n'mnkr) \equiv (ijkrm')(i''n')(j''mnkr)$, and this latter can be obtained from two A factors and appears in the consideration of the Q-factors.

We next consider type $B = (ijkrm')(in' \mid Y)$. If one of the component variables of Y is convolved to from u, we have the types, $(ijkrm')(n'abcd)i'$, $(ijkrm')(i'n'abc)$, $(ijkrm')(i'n'a''f''s)(abcde'')$. Of these, the first two are obviously reducible to simpler types and the last, formed from two B factors is also reducible. For $a \neq i, m$, or n and so $a = j, k$ or r and

$$(ijkrm')(in'af''s)(abcde'') \\ \equiv (ijkrf'')(inams)(abcde'') + (ijkrs)(inaf''m)(abcde'').$$

Both of the last two terms have a simple bracket factor as an actual factor and hence this type is also reducible.

Since, in the formation of new u-factors, A and D can only occur with simple factors of the type (ij), (ijk) etc., it is easily seen that A and D do not give rise to any new u-factors. Thus there are no new u-factors.

New p-factors. Type C cannot occur with a p-factor, for *

$$(ijkr'm)(ijn's) \equiv (ijkrs)(ijnm) + (ijkr)(ijnms).$$

Thus C yields only six factors of the type $(ijkrm'u)(ijn'P)$. Further type B cannot occur with a p-factor, for $(ijkrm')(in'a''b''')$ is reducible,* where a and b appear from factors of types A, D, F, G, H or I. Also $(ijkrm')(in'ab)$ is reducible, since a, b must be two of r, j, k. We are left to consider $(ijkrm')(in'a''b'')(c''def)$ arising from a B and an H factor. But this is impossible, since i cannot be one of a, b, c or one of d, e, f, since a, b, c and d, e, f are interchangeable. If however one of the components of u occur in a three-factor, we have the type $(ijkrm')(n'abc)i'$, where abc cannot contain i or both of m, n. If $a = m$, we have $(ijkrm')(n'mkr)i' \equiv (ijkrm)(n'm'kr)i'$, and this latter is reducible, being the product of $(ijkrm)$ and $i'(n'm'kr)$. Hence $abc = jkr$ and $(ijkrm')(n'jkr)i' \equiv (ijkr)(n'm'jkr)i'$, and the latter is a product of two simpler factor types.

Type $F = (ijkr')(m'n')$ cannot occur with a p-factor, for as before $(ijkr')(m'n'ab)$ and $(ijkr')(m'n'a''b)$ are both reducible and

$$(ijkr')(m'n'a''b'')(c''def),$$

* *Loc. cit.*, p. 408, formula (17).

from two F factors, is impossible, since two of a, b, c cannot be the same as two of r, m, n or two i, j, k. Hence we must consider $(ijkr')(n'abc)m'$. In this a, b, c cannot contain two of r, m, n and must therefore contain two of i, j, k which is impossible.

Type G cannot appear with a p-factor, for $(ijkr')(im's''a''')$ and $(ijkr')(im's''a)$ are reducible.* But in $(ijkr')(im'ae'')(abcd'')$, formed from two G factors, a is distinct from i, r, j, m and k and must therefore be equal to n. Similarly i is not equal to any of a, b, c, d, e. Accordingly we must consider $(ijkr')(im'nm'')(nr''jk) \equiv (ijkr')(imnr)(nm'jk)$, which is reducible, and

$$(ijkr')(im'nr'')(nk''mj) \equiv (ijkn)(imrr'')(nk''mj) + (ijkr'')(imnr)(nk''mj),$$

and this is reducible, unless n, r, k are successive integers. Similarly i, j, k; i, m, r; n, m, j must all be sets of successive integers and on trial this is found to be impossible. We have still to consider the type $(ijkr')(m'abc)i'$, where a, b, c cannot contain i or both of r, m or both of j, k, since r, m and j, k are interchangeable in G. Hence we have $(ijkr')(m'njr)i'$ and this type is reducible, if i, m, n or i, j, k are not successive integers. But

$$(ijkr')(m'njr)i' \equiv (ijkn'')(m'r'jr'')i'$$

and this latter is reducible unless n, r are successive integers. Similarly j, r must be successive integers. Since the only possible type of G factor now is $(3124')(35')$, $n = 6$ and accordingly n, j, r cannot be successive integers.

Since factors of the type A, F, H and I can only appear with factors of the types (ij), (ijk) or with factors in which the symbols are not convolved, it is easy to show that no new p-factors arise from considering them. We take type H as an illustration. In the type $(ijkr')(m'abc)$ a, b, c cannot include r or m and must be i, j, k or i, j, n. But $(ijkr')(m'ijk)$ is reducible and $(ijkr')(m'ijn) \equiv (ijkn)(mrij)$, where n is still convolved with i and j. The only other possibilities are $(ijkr')(m'abs'')$ and $(ijkr')(m'ab''c''')$ both of which are obviously reducible. Accordingly there are no new p-factors.

New P-factors. The types F and G cannot occur with a P-factor, for †

$$(ijkr')(im's) \equiv (ijks)(imr) + (ijk)(imsr)$$

and no convolutions of successive symbols have been disturbed. A similar

* *Loc. cit.*, formula (17).

† *Loc. cit.*, formula (17).

proof holds for the type F. Accordingly F yields only three factors of the type $(ijkr'p)(m'n'Q)$ and G twelve of the type $(ijkr'p)(im'Q)$.

If type B occur with a further P-factor, it may occur with a single symbol thus yielding $(ijkrm')(in'a)$, where a is not equal to $i, m,$ or n and so this reduces to type C. But B may occur with a simple two-factor giving the type $(ijkrm')(in'a'')b''$, where neither of a, b is equal to i or one of $n, m,$ since $(ijkrm')(in'a'')b'' \equiv (ij'k')(r'imna'')b''$. Therefore we may take $(ijkrm')(in'j'')k''$ as typical. But this reduces to $(ijkrm')(kn'j)i$, unless i, m, n are successive integers. From the list of B factors we have only four possible types, $(3'1456)(12'4'')5''$, $(3'1456)(12'5'')6''$, $(34562')(31'4'')5''$, $(34562')(31'5'')6''$. Of these the first is equivalent to

$$
\begin{aligned}
&(14''5'')(1236''4')5' \\
&\quad \equiv (14''5'')(12345)6'' + (14''5'')(1'''2'''5'6''4')3''' \equiv (145)(1'2'564)3',
\end{aligned}
$$

and this last is a product of two simpler factors; similarily the second is reducible; the third is equivalent to

$$(34'5')(3126'4'')5'' \equiv (345)(4561'2')3',$$

and so is reducible; the fourth is equivalent to

$$(34'5')(3125''6')6'' \equiv (356)(1'2'456)3'$$

but is not reducible, since 3, 4, which was originally convolved, is no longer convolved. Thus we have the new factor type $(34562')(31'5'')6''$ and the factor similar to it. This type need not be considered farther, for, if the extra variable attached to $6''$ appear in a P-factor, the resulting factor type obviously reduces except in the case $(34562')(31'5'')(6''ij)$, where i, j are successive integers. But neither of i, j can be 5 or 3 and so ij must be 12 and $(34562')(31'5'')(6''12) \equiv (34562')(312)(561')$, which is reducible. Further $(315')(6'i) \equiv (31i)(65) + (31)(65i)$, and so, if the extra variable is convolved to form Q, no new factor type is obtained.

If a new P-factor is formed from two B factors, we have the general type $(ijkrm')(abcde'')(in'a'')f''$. This type is reducible, unless i, n, m and a, e, f are both sets of successive integers and also if these two sets coincide. From the list of B factors we see that we must consider

$$
\begin{gathered}
(12'6'')(14563')(63214'')5'', \quad (12'5'')(43216'')(14563')4'', \\
(31'4'')(34562')(43216'')5''.
\end{gathered}
$$

Of these the first is equivalent to

$$(12'6)(14563')(6''3214'')5''+(12'3''')(14563')(66''2'''1'''4'')5'',$$
$$\equiv(12'6)(14563')(123,456)+(123)(14563')(6542'1')6,$$
$$\equiv(12'6)(14563')(123,456);$$

similarly the second and the third are reducible.

We now consider the possibility of a new factor type formed by a B factor and one other factor of the types A, D etc. in turn.

Types A and B. $B=(ijkrm')(in')$, $A=(abcde')f'$. In such a factor e, f is not the same as m, n nor is one of e, f equal to i. If $e=m$, we have

$$(ijkrm')(ijknm'')(r''in')$$
$$\equiv(r''imnj')(ik'r')(ijknm'')\equiv(rimnj')(irk')(ijknm),$$

and the last of these is reducible. Hence $e, f=j, k$ and we must consider $(ijkrm')(inmrj'')(in'k'')$. But there is only one type of A factor and so, from the list of B factors, we are left with

$$(2'3456)(1'65'')(6''3214'')\equiv(2'3456)(1'4562'')(63''1''),$$
$$\equiv(2''3456)(14562)(63''1'')\equiv(2''3456)(14562)(631''),$$

and

$$(2'3456)(43215'')(46''1')$$
$$\equiv(2'3456)(43''2'')(4651'1'')\equiv(1''3456)(432'')(46512),$$

both of which are reducible.

Types B and D. $B=(ijkrm')(in')$, $D=(abcd'e')f'$. Since abc and def are interchangeable in D, i may be taken equal to e and we have the type $(ijkrm')(in'f'')(abcid'')$, which is the same as a factor arising from one A and one B factor and accordingly has already been considered.

Types B and H. $B=(ijkrm')(in')$, $H=(cdeb')a'$. We have the factor type $(ijkrm')(in'a'')(b''cde)$, which reduces as before, if either of a, b is one of i, m, n. Farther, since $(ina')(b'cde)\equiv(i'ba)(n'cde)+(inab)(cde)$, this type reduces, unless i, m, n are consecutive integers and also if i is equal to one of c, d, e. Accordingly $i=f$, and we have

$$(fabcd')(fe'a'')(b''cde)\equiv(fabcd')(be'a)(fcde)+(fabcd')(fe'ab)(cde),$$

and both terms on the right are products of simpler factor types.

Types B and I. $B=(ijkrm')(in')$, $I=(abc'e'')d'f''$. This type is not possible, since ab, cd, and ef are interchangeable in I and neither of c, d can equal i in B.

Types B and J. $B=(ijkrm')(in')$, $J=(bcd')a'$. As in previous cases i, n, m must be consecutive integers and none of a, b, c, d can belong to i, n, m and this is impossible.

P-factors involving type A but not type B. If A occurs with an ordinary two-factor, the resulting factor is obviously reducible. The only other possibility is

$$(ijkrm')(n'a''b''')\equiv(ijkra'')(mnb''')$$
$$+(ijkrb''')(mna'')+(i'j'k')(r'mna''b''')\equiv(i'j'k')(r'mna''b''').$$

Accordingly, if a, b arise from two other A factors, this type is reducible, since there are only two distinct A factors. Thus there is no new P-factor formed by three A factors. Further, since in A $mn=12$ or 56, the combination of an A factor an F factor and any other factor cannot occur. Moreover, since

$$(ijkr')(m'ts)\equiv(ijk)(rmts)+(ijkt)(rms)+(ijks)(mtr),$$

any factor formed from factors of types A, H, I or J is reducible. Similarly type D cannot appear with types H, I or J. If D occur with an ordinary two-factor, we have $(ijkr'm')(n'mj)$, and this is of type C.

We must now consider the types H, I and J with ordinary simple factors (ab). In type $H=(ijkr')(m'ab)$ neither of a, b is equal to one of r, m and so we have $(ijkr')(m'ab)=(ijk)(rmab)+(ijka'')(rmb'')$ and this is reducible, since at least one of a, b is the same as one of i, j. The factor $I=(ijk'm'')r'n''$ with the simple factor (ab) is impossible, since neither of k, r in I can equal one of a, b and since k, r; m, n; i, j are interchangeable in I. But J and the simple factor (ab) yield the new type $(ijk')(m'ab)$ $=(123')(4'56)$.

The only type, which we have neglected, is $(ijk'm'')(r'n''a)$, and this is reducible, for

$$(ijk'm'')(r'n''a)\equiv(ijkr)(mna)+(ijk'a)(r'mn)+(ijk')(mnr'a),$$

and each term on the right has at least one less broken convolution.

New Q-factors. Type J cannot occur with a new Q-factor, for $(ijk')(m'a)$ $\equiv(ija)(km)+(ij)(mka)$. Two factors of the type A yield the new factor type $(12345')(6'1'')(2''3456)$. The only possibility from one A factor and one D factor is

$$\begin{aligned}(12345')(6'1'')(2''3''456) &\equiv (12345')(6'6'')(1234''5''),\\ &\equiv (1''2'')(3''4''566')(1234'5') \equiv (1''2'')(3''4''456)(12356),\\ &\equiv (1'4)(3'2'456)(12356),\end{aligned}$$

and this last is reducible. From A and H we have the type

$$(ijkrm')(n'a'')(b''cde) \equiv (ijkrm')(n'abc'')(d''e'') \equiv (ijkra')(mnb'c'')(d''e''),$$

where we have neglected terms, which are reducible. From the first identity we see that neither of n, m can be the same as one of a, b and from the second, that neither of a, b can be the same as one of i, j, k, r. But this is impossible and so the type reduces. Since in I i, j; k, r; m, n are interchangeable, no new factor arises from A and I. From two D factors we have the type

$$(1234'5')(6'6'')(1234''5'') \equiv (1234''5'')(1236'' \mid x'y'z'w')(456 \mid xyt') \equiv 0$$

by the convolution of 4, 5, 6 in the first factor. From D and H we have the possibility $(ijkr'm')(n'a'')(b''cde) \equiv (ijkr'm')(n'abc'')(d''e'')$. This latter is reducible, if a and b appear among r, m, n or among i, j, k. Accordingly $a, b = 3, 4$ and cde is of the type 126 or 125. Therefore a, b and c, d can be interchanged and since $cd = 12$, this new factor reduces. Similarly the combinations of D with I, and H with I are reducible. From two H factors we have the type

$$\begin{aligned}(ijkr')(m'a'')(b''cde) &\equiv (ijkr')(m'abc'')(d''e'')\\ &\equiv (ijkc'')(rmab)(d''e'') + (ijka')(rmb'c'')(d''e'').\end{aligned}$$

From the first identity we see that neither of a, b is the same as one of r, m and that therefore one of a, b must equal one of i, j, k. Accordingly this type reduces by the second identity. No new type arises from two I factors, but the single I factor yields the new type $(123'5'')(4'6'')$.

We have now found all possible cases, in which a convolution of successive symbols has been disturbed. These new factor types, together with the 63 simple factors form the prepared system. In § 1 these factors are listed and defined in terms of the symbols a, r etc. of the two quadratics, and if the factors I are removed from this list, we are left with a prepared system similar to that used by Turnbull in his paper on Two Quaternary Quadratic Forms.*

THE JOHNS HOPKINS UNIVERSITY.

---

* H. W. Turnbull, "The Simultaneous System of Two Quadratic Quaternary Forms," *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 18 (1917), parts 1 and 2, pp. 70-94.

A VARIETY REPRESENTING PAIRS OF POINTS OF SPACE.

By F. R. Sharpe.

1. *Introduction.* If $X \equiv (x_1, x_2, x_3, x_4)$ and $Y \equiv (y_1, y_2, y_3, y_4)$ are any two points in space, the 10 quantities,

$$g_{ij} = x_i y_j + x_j y_i, \qquad (i, j = 1, 2, 3, 4), \tag{1}$$

may be taken as coördinates of a point P in S_9. The locus of P, when X and Y, vary is a six-dimensional variety V_6. It will be shown that V_6 is rational, being mappable on the S_6, $(g_{41}, g_{42}, g_{43}, g_{44}, g_{23}, g_{31}, g_{12})$. When X and Y are corresponding points of an involutional transformation, the locus of P is a three-dimensional variety. The case of the cubic inversion $y_i = 1/x_i$ has been partially discussed by Emch,* but the reasons given for rationality are insufficient. The general cubic involution will be shown to be rational by mapping it on the S_3, $(g_{41}, g_{42}, g_{43}, g_{44})$. The proof differs essentially from and is much simpler than the proof given by Sharpe and Snyder.† It is also shown that the complex of lines joining corresponding points of the general cubic involution can also be mapped on S_3, $(g_{41}, g_{42}, g_{43}, g_{44})$ and its equation is obtained.

2. *The variety* V_6. If $p_{ij} = x_i y_j - x_j y_i$ are line coördinates of XY, they are connected with the g_{ij}, by the identities,

$$p_{ij} p_{kl} = g_{il} g_{jk} - g_{ik} g_{jl}. \tag{2}$$

Hence

$$p_{42} p_{43} = g_{42} g_{43} - g_{44} g_{23} = g_1, \qquad p_{43} p_{41} = g_{43} g_{41} - g_{44} g_{31} = g_2, \tag{3}$$
$$p_{41} p_{42} = g_{41} g_{42} - g_{44} g_{12} = g_3$$

so that

$$p_{41} = (g_2 g_3 / g_1)^{1/2}, \quad p_{42} = (g_3 g_1 / g_2)^{1/2}, \quad p_{43} = (g_1 g_2 / g_3)^{1/2}, \tag{4}$$

where the signs of the radicals are all $+$ or all $-$. We have therefore

$$2y_4 x_4 = g_{44}, \qquad 2y_4 x_1 = g_{41} - (g_2 g_3 / g_1)^{1/2}, \tag{5}$$
$$2y_4 x_2 = g_{42} - (g_3 g_1 / g_2)^{1/2}, \qquad 2y_4 x_3 = g_{43} - (g_1 g_2 / g_3)^{1/2},$$

and hence

---

* A. Emch, *American Journal of Mathematics*, Vol. 41 (1926), pp. 21-44.

† F. R. Sharpe and Virgil Snyder, *Transactions of the American Mathematical Society*, Vol. 25 (1923), pp. 1-12.

(6) $$g_{11}g_{44} = g^2{}_{41} - g_2g_3/g_1, \qquad g_{22}g_{44} = g^2{}_{42} - g_3g_1/g_2,$$
$$g_{33}g_{44} = g^2{}_{43} - g_1g_2/g_3.$$

It follows from (6) that V_6 is rational, being mapped by (6) on the $S_6(g_{41}, g_{42}, g_{43}, g_{44}, g_{23}, g_{31}, g_{12})$. The equations (5) express the (1, 2) correspondence between the spaces $S(x, y)$ and S_6.

3. *The general cubic involution I between X and Y.* The involution is defined by three equations bilinear in (x) and (y), which, by choosing as vertices of the tetrahedron of reference four of the eight invariant points, can be taken in the form.*

(7) $$g_{23} = a_1g_{41} + a_2g_{42} + a_3g_{43}, \qquad g_{31} = b_1g_{41} + b_2g_{42} + b_3g_{43},$$
$$g_{12} = c_1g_{41} + c_2g_{42} + c_3g_{43}.$$

When account is taken of (7), the equations (5) express the (1, 2) correspondence between the spaces $S(x)$ and $S_3(g_{41}, g_{42}, g_{43}, g_{44})$.

Hence I is rational. From (5) it can be seen that the image of a plane in $S(x)$

(8) $$A_1x_1 + A_2x_2 + A_3x_3 + A_4x_4 = 0$$

is the surface in S_3

(9) $$(A_1g_{41} + A_2g_{42} + A_3g_{43} + A_4g_{44})^2g_1g_2g_3 = (A_1g_2g_3 + A_2g_3g_1 + A_3g_1g_2)^2,$$

which is apparently of order 8 in the g_{ij}. The terms independent of g_{44} however, vanish identically, so that g_{44} is a factor and the image of a plane in $S(x)$ is a surface F_7 of order 7 in S_3. The image of F_7 is the original plane (8) and the cubic surface

(10) $$A_1y_1 + A_2y_2 + A_3y_3 + A_4y_4 = 0,$$

the y_i being found from (8) by substituting from (1) and solving in terms of the x_i.

The surfaces

(11) $$C_1g_{41} + C_2g_{42} + C_3g_{43} + C_4g_{44} = 0,$$

images of the planes in S_3 are quartic surfaces through the sextic curve common to the cubic surfaces $y_i = 0$ and through the three straight lines, $x_4 = 0$, and one of $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, in which $x_4 = 0$ meets $y_4 = 0$. Any two of the surfaces (11) meet in a residual curve of order 7, image of a line in S_3. It can be shown from (9) that the images of the three invariant points (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0) in $S(x)$ are the three quadrics $g_1 = 0$, $g_2 = 0$, $g_3 = 0$ in S_3.

The three quadrics meet in the vertices of the tetrahedron of reference

* Compare F. R. Sharpe and Virgil Snyder, *Transactions of the American Mathematical Society*, Vol. 25 (1923), pp. 1-12.

and in four other points. These four points and $(0, 0, 0, 1)$ in S_3 are the images of the four invariant points which are not vertices in $S(x)$ and of $(0, 0, 0, 1)$.

4. *The complex of lines X Y.* From (2) we have, in addition to (3), the relations

$$(12) \quad p_{41}p_{23} = g_{43}g_{12} - g_{42}g_{31} = g_4, \qquad p_{42}p_{31} = g_{41}g_{23} - g_{43}g_{12} = g_5,$$
$$p_{43}p_{12} = g_{42}g_{31} - g_{41}g_{23} = g_6,$$

and therefore the quadratic identity

$$p_{41}p_{23} + p_{42}p_{31} + p_{43}p_{12} = 0.$$

From (4) and (12) we find

$$(13) \quad p_{41}/g_2g_3 = p_{42}/g_3g_1 = p_{43}/g_1g_2$$
$$= p_{23}/g_4g_1 = p_{31}/g_5g_2 = p_{12}/g_6g_3 = (g_1g_2g_3)^{-\frac{1}{2}}.$$

The p_{ij} and g_{ij} are connected by the identities

$$(14) \qquad g_{ij}p_{kl} + g_{jk}p_{li} + g_{jl}p_{ik} = 0.$$

Hence

$$(15) \quad g_{12}p_{34} + g_{23}p_{41} + g_{42}p_{13} = 0, \quad g_{12}p_{43} + g_{41}p_{32} + g_{31}p_{24} = 0,$$
$$g_{31}p_{42} + g_{43}p_{21} + g_{23}p_{14} = 0, \quad g_{41}p_{23} + g_{42}p_{31} + g_{43}p_{12} = 0.$$

Any three of these equations are linearly independent, but the sum of the four vanishes identically.

From (14) we also have

$$(16) \quad g_{44}p_{23} + g_{42}p_{34} + g_{43}p_{42} = 0, \qquad g_{44}p_{31} + g_{43}p_{14} + g_{41}p_{43} = 0,$$
$$g_{44}p_{12} + g_{41}p_{24} + g_{42}p_{41} = 0.$$

When $g_{41}, g_{42}, g_{43}, g_{44}$ are given, then (7) and (13) determine the p_{ij} which satisfy (15) and (16). If we eliminate the g_{ij} from (7) and (15) we have the equation of the cubic complex.

$$(17) \quad \begin{vmatrix} a_1 & a_2 & a_3 & -1 & 0 & 0 \\ b_1 & b_2 & 3 & 0 & -1 & 0 \\ c_1 & c_2 & c_3 & 0 & 0 & -1 \\ 0 & p_{13} & 0 & p_{41} & 0 & -p_{43} \\ p_{32} & 0 & 0 & 0 & -p_{42} & p_{43} \\ 0 & 0 & p_{21} & -p_{41} & p_{42} & 0 \end{vmatrix} = 0.$$

Conversely if the p_{ij} satisfy (17) and the quadratic identity, then (7), (15), and (16) determine the g_{ij}. Hence the complex is mappable on S_3. The cubic inversion is the special case when $a_1 = b_2 = c_3 = 1$, the other coefficients in (7) being zero.

CORNELL UNIVERSITY.

ON SEMI-METRIC SPACES.*

By Wallace Alvin Wilson.

1. Let Z be a set of points to each pair of which corresponds a positive real number called the distance between them. If a and b are any two points, we designate this distance by ab, and postulate that the following axioms are satisfied:

I. $ab = ba$.

II. $ab = 0$ *if and only if* $a = b$.

A space which satisfies these conditions and in which limiting points are defined in the usual way is called by Frechet an E-space and by Menger a *semi-metric* space.

As a semi-metric space becomes metric when the so-called triangle axiom is added, it is natural to classify these spaces by the degree to which the triangle axiom is approximated. Hence we are led to the following additional axioms.

III. *For each pair of points a and b there is a positive number r such that for every point c, $ac + bc \geqq r$.*

IV. *For each point a and each positive number k there is a positive number r such that, if b is a point for which $ab \geqq k$ and c is any point, $ac + bc \geqq r$.*

V. *For each positive number k there is a positive number r such that, if a and b are any points for which $ab \geqq k$ and c is any point, $ac + bc \geqq r$.*

If Axiom V is further strengthened by requiring r to equal k, our space becomes metric. Furthermore, E. W. Chittenden † has shown by an equivalent definition that a semi-metric space in which Axiom V is valid is homeomorphic with a metric space. It is the purpose of this article to supplement Chittenden's work by investigating this question for the weaker Axiom IV and also to discuss certain other properties of these spaces.

2. Examples of spaces consisting of enumerable sets of points can be

* Presented to the American Mathematical Society, February, 1931.

† "On the equivalence of ecart and voisinage," *Transactions of the American Mathematical Society*, Vol. 18, pp. 161-166.

constructed, which show that Axiom V is effectively stronger than Axiom IV, and that Axiom IV is effectively stronger than Axiom III.

In demonstration it is often convenient to make use of the following easily proved properties. *For Axiom III to be valid it is necessary and sufficient that there do not exist two points a and b and a sequence $\{c_i\}$ such that $ac_i + bc_i \to 0$. For Axiom IV to be valid it is necessary and sufficient that there do not exist a point a and two sequences $\{b_i\}$ and $\{c_i\}$ such that $ac_i + b_ic_i \to 0$ but not $ab_i \to 0$. For Axiom V to be valid it is necessary and sufficient that there do not exist three sequences $\{a_i\}$, $\{b_i\}$, and $\{c_i\}$, such that $a_ic_i + b_ic_i \to 0$, but not $a_ib_i \to 0$.*

In the case of each of the three axioms it is readily shown that there is a greatest r for which $ac + bc \geqq r$. In future reference to the axioms it will be understood that r represents this greatest value. In Axiom III, r is a function of a and b; in Axiom IV, $r = f(a, k)$; and in Axiom V, $r = f(k)$. In Axioms IV and V, r is a monotone ~~increasing~~ non-decreasing function of k. Since c may coincide with a or b, it is clear that ~~in the last two cases $r \leqq k$~~ and in the first case $r \leqq ab$. Finally, in the last two cases the inequality $ac + bc < r$ implies that $ab < k$.

3. The scope of our work is limited by the following theorems:

I. *For Axiom III to be valid in a semi-metric space it is necessary and sufficient that no sequence converge to more than one limit.*

II. *In a semi-metric space satisfying Axiom IV every derived set is closed.*

The proof of the first of these theorems is self-evident. To prove the second let $A = \{a\}$ be any set, $A' = \{b\}$ be the derived set of A, and c be a limiting point of A'. If c does not lie in A', there is a $k > 0$ such that for no point a in $A - c$ is $ca < k$.

Since Axiom IV is valid, let $r = f(c, k)$; then for every point x, $cx + ax \geqq r$. Take $e < r/2$. Since c is a limiting point of A', there is some point b for which $cb < e$. Since b lies in A', it is a limiting point of A, and also of $A - c$. Hence there is a point a in $A - c$ for which $ba < e$. Then $cb + ab < r$, contrary to the statement at the beginning of the paragraph. Hence c lies in A' and A' is closed.

COROLLARY. *In a semi-metric space satisfying Axiom IV the set of inner points of any set is a region.*

It can be shown by an example that the converse of Theorem II is not valid, and that the theorem itself is not valid in general if only Axiom III

holds. Consequently it does not seem profitable to give any attention to spaces satisfying Axiom III only.

4. The simpler theorems regarding closed sets and regions in general metric spaces may now be proved valid for semi-metric spaces satisfying Axiom IV in the usual way. In working with these concepts, however, it is necessary to keep in mind two respects in which these spaces differ radically from metric spaces. These are:

In a semi-metric space satisfying Axiom IV the Cauchy criterion for the convergence of a sequence of points to a point is not necessary, and the distance function may have discontinuities.

To see this consider a space Z consisting of a point a and an enumerable set of points $\{a_i\}$, where the distances are defined as follows: $aa_i = 1/i$ and $a_ia_j = 1$ if $i \neq j$. It is easily seen that Axiom IV is valid. Obviously $a_i \rightarrow a$, but the Cauchy property is not valid. Finally, as $a_2a = 1/2$ and $a_2a_i = 1$ for every $i \geqq 2$, it is evident that a_2a_i does not converge to a_2a, and so the distance function is not continuous.

It is convenient to call the set of points $\{x\}$ whose distances from a fixed point a are less than some fixed k a sphere, but on account of the possible discontinuity shown above a sphere may fail to be a region; also the sets for which $ax \leqq k$ or $ax \geqq k$ may fail to be closed sets. However, if S is a sphere of center a and radius k, and $r = f(a, k)$, it readily follows from Axiom IV that every point of the sphere of center a and radius r is an inner point of S. Likewise, if T denotes the set $\{x\}$ for which $ax \geqq k$, $\bar{T}$ contains no point y for which $ay < r$.

Furthermore, if s' is a sphere of radius $r' < r$ and center a, $\bar{s}' \subset S$. Hence, if $r' = f(a, r)$ and $r'' < r'$, the sphere s of center a and radius r'' is such that $\bar{s} \cdot \bar{T} = 0$. Such a sphere may be called an *inner sphere* corresponding to a and k.

5. Before proceeding further, we turn to a brief consideration of semi-metric spaces in which Axiom V is valid. We first note that in such a space, *every convergent sequence satisfies the Cauchy convergence criterion.* For otherwise we would have a sequence $\{a_i\}$, where $a_i \rightarrow a$, and a constant $k > 0$ such that for every integer i' there would exist an i and a j greater than i' for which $a_ia_j \geqq k$. But, as $i' \rightarrow \infty$, $a_ia + a_ja \rightarrow 0$, which contradicts Axiom V.

Now let Z be any semi-metric space and a, b, and c be any three points. Also let $g(e)$ be a positive function of the positive real variable e which

converges to zero with e, and let the relations $ac < e$ and $bc < e$ always give $ab < g(e)$. Such a space Frechet calls a space with a *regular ecart.*

THEOREM I. *A semi-metric space in which Axiom V is valid has a regular ecart, and conversely.*

Proof. If Axiom V is valid, there is a monotone increasing function $r = f(k)$ such that $ab \geqq k$ implies that $ac + bc \geqq r$ for any three points of the space. If k has a lower bound $k' > 0$, r has a lower bound $r' > 0$; in this event let us define $r = f(k)$ for $0 \leqq k \leqq k'$ by the relation $r/k = r'/k'$. Now let $h = h(r)$ be 4/3 the lower bound of all values of k for which $r = f(k)$. Then h is one-valued and converges to zero as $r \rightarrow 0$. If $ac < r/3$ and $bc < r/3$, then $ab < h$. Taking $e = r/3$, $h = g(e) = h(r)$ is the required function.

To show the converse, assume that Axiom V fails. Then there would be a $k > 0$ and sequences $\{a_i\}$, $\{b_i\}$, and $\{c_i\}$ such that $a_ib_i \geqq k$ for every i, but $a_ic_i + b_ic_i \rightarrow 0$. Taking e so that $g(e) < k$, there is an i such that $a_ic_i < e$ and $b_ic_i < e$, while $a_ib_i \geqq k > g(e)$, contrary to the definition of regular ecart.

The above theorem together with Chittenden's theorem (*loc. cit.* in § 1) gives the following result.

THEOREM II. *For a semi-metric space Z to be uniformly homeomorphic with a metric space it is necessary and sufficient that Axiom V be valid.*

That the condition is sufficient is proved by Chittenden. It remains to prove that it is necessary.

Let $Z = \{x\}$ and let $M = \{y\}$ be metric and uniformly homeomorphic with Z. If Axiom V is not valid in Z, there is a $k > 0$ and three sequences $\{x_i\}$, $\{x_i'\}$, and $\{x_i''\}$ such that $x_i'x_i'' \geqq k$ and $x_ix_i' + x_ix_i'' \rightarrow 0$. Let the corresponding sequences in M be $\{y_i\}$, $\{y_i'\}$, and $\{y_i''\}$. Since the homeomorphism is uniform, $y_iy_i' + y_iy_i'' \rightarrow 0$ and consequently, by the triangle axiom, $y_i'y_i'' \rightarrow 0$. But again this result requires that $x_i'x_i'' \rightarrow 0$, which is a contradiction.

6. THEOREM. *Let Z be a semi-metric space satisfying Axiom IV. Then Z is homeomorphic with a semi-metric space satisfying Axiom V.*

*Proof.** For each point x of Z define two descending sequences $\{r_i(x)\}$ and $\{s_i(x)\}$ as follows:

---

* The following proof is based upon one by Alexandroff and Urysohn, "Une condition nécessaire et suffisante pour qu'une classe L soit une classe D," *Comptes Rendus des Séances de l'Académie des Sciences*, Vol. 177, pp. 1274-1276.

$$\begin{array}{ll} s_0(x) = 1, & r_1(x) = \tfrac{1}{2}f(x,1), \\ r_2(x) = \tfrac{1}{2}f[x, r_1(x)], & s_1(x) = \tfrac{1}{2}r_2(x), \cdots, \\ s_{i-1}(x) = \tfrac{1}{2}r_{2(i-1)}(x), & r_{2i-1}(x) = \tfrac{1}{2}f[x, s_{i-1}(x)], \\ r_{2i}(x) = \tfrac{1}{2}f[x, r_{2i-1}(x)], & s_i(x) = \tfrac{1}{2}r_{2i}(x), \cdots. \end{array}$$

Here f is the function of Axiom IV. It is clear that $s_i(x) \leqq \tfrac{1}{2}^{3i}$ for every i and x. For each i and x let $U_i(x)$ be the sphere of center x and radius $s_i(x)$ and K_i be the set of spheres $\{U_i(x)\}$ as x ranges over Z.

We now prove that, if $D_i = U_i(x) \cdot U_i(y) \neq 0$ for some $i \geqq 1$, then $U_i(x) + U_i(y)$ is contained in either $U_{i-1}(x)$ or $U_{i-1}(y)$. Let u be a point of D_i. There are two cases: $s_i(x) \geqq s_i(y)$ or $s_i(x) < s_i(y)$.

In the first case $xu + uy < 2s_i(x) = r_{2i}(x)$, whence $xy < r_{2i-1}(x)$ by Axiom IV. If z lies in $U_i(y)$, $yz < s_i(y) \leqq s_i(x)$. Hence $xy + yz < 2r_{2i-1}(x) = f[x, s_{i-1}(x)]$ and so $xz < s_{i-1}(x)$ by Axiom IV. Obviously $xz < s_{i-1}(x)$ if z lies in $U_i(x)$. Hence $U_i(x) + U_i(y) \subset U_{i-1}(x)$ in this case.

In the second case $xz < s_i(x) < s_i(y)$ for any z in $U_i(x)$; whence $xy + xz < 2r_{2i-1}(y)$ and $yz < s_{i-1}(y)$. Obviously $yz < s_{i-1}(y)$ if z lies in $U_i(y)$. Thus in this case $U_i(x) + U_i(y) \subset U_{i-1}(y)$.

Now let us define a new distance $d(x,y)$ for each pair of points of Z as follows and call the new space Z'. If no sphere of any K_i contains $x + y$, let $d(x,y) = 1$. If no sphere of K_{i+1} contains $x + y$, but some sphere of K_i contains $x + y$, let $d(x,y) = \tfrac{1}{2}^i$.

We first show that Z' is semi-metric. If $x \neq y$, $d(x,y)$ exists and is positive unless some sphere of every K_i contains $x + y$. Then there would be a sequence $\{U_i(c_i)\}$ of spheres whose respective centers and radii are $\{c_i\}$ and $\{s_i(c_i)\}$, each of which would contain $x + y$. Since $s_i(c_i) \to 0$, this gives the contradiction that $c_i \to x$ and $c_i \to y$ in Z.

If x, y, and z are three points, $d(x,z) = 1/2^i$, and $d(y,z) = 1/2^j$, $j \geqq i$, we have two spheres $U_i(a)$ and $U_i(b)$ of K_i containing $x + z$ and $y + z$, respectively. Then by the above, these spheres are contained in either $U_{i-1}(a)$ or $U_{i-1}(b)$. In both cases $d(x,z) \leqq 1/2^{i-1}$, and so the distance $d(x,y)$ is a regular ecart. Hence Axiom V is valid in Z' by § 5, Theorem I.

Let a be a fixed point, $k > 0$, and S be the set of points $\{x\}$ for which $d(a,x) < k$. Take i so large that $1/2^i < k$. Then $U_i(a) \subset S$. For, if x lies in $U_i(a)$, $a + x$ lies in $U_j(a)$ for $0 \leqq j \leqq i$ and hence $d(a,x) \leqq 1/2^i$. That is, S contains every point x for which $ax < s_i(a)$.

Now let S be the set of points $\{x\}$ for which $ax < k$. Take i so large that $s_{i-1}(a) < k$ and j so large that $2/2^{3j} < s_i(a)$. Let S' be the set of points $\{x\}$ for which $d(a,x) < 1/2^j$. Then for each x in S', $a + x$ lies in

some $U_j(b)$; i. e., $ab + bx < 2s_j(b) < 2/2^{3j} < s_i(a)$, whence $ax < s_{i-1}(a) < k$. That is, S contains every point x for which $d(a, x) < 1/2^j$.

The last two paragraphs prove that Z and Z' are homeomorphic.

COROLLARY. *Let Z be a semi-metric space satisfying Axiom IV. Then Z is homeomorphic with a metric space.*

This is an immediate consequence of the above theorem and Chittenden's theorem (§ 5, Theorem II). Note, however, that in this case homeomorphism between a semi-metric and a metric space does not imply that Axiom IV is valid in the former space.

7. A semi-metric space Z which satisfies Axiom IV and contains an enumerable set $E = \{a_i\}$ such that every point of Z is the limit of a subsequence chosen from E will be called *separable,* as usual. Following Frechet, we shall call Z perfectly separable or *p-separable,* if for each $k > 0$ each point x of Z lies in an inner sphere corresponding to k and some a_i. (See § 4.)

The author does not know whether separability implies p-separability or not. It is possible to construct a space Z containing an enumerable set E dense in Z and such that for a given $k > 0$, Z is not covered by a set of inner spheres having their respective centers in E and corresponding to k, but this does not show that there is no enumerable set having the desired property. Since by the previous section a separable semi-metric space satisfying Axiom IV is homeomorphic with a separable metric space and a separable metric space is always p-separable, it might appear that separability implies p-separability in semi-metric space also. This does not follow, however, because the homeomorphism between the two spaces may fail to be uniform.

It is easy to show that a p-separable semi-metric space satisfying Axiom IV is homeomorphic with a separable semi-metric space without using Chittenden's theorem. In brief we first prove the theorem that, if $\bar{A} \cdot B + A \cdot \bar{B} = 0$, then there are disjoint regions R and S containing A and B, respectively, in much the same way as for metric spaces. Then Urysohn's proof * that for two disjoint closed sets A and B, there is a continuous function $f(x)$ such that $f(x) = 0$ in A, $f(x) = 1$ in B, and $0 \leqq f(x) \leqq 1$ in $Z - (A + B)$ is applicable. Then, for each k_j of a descending sequence $\{k_j\}$ converging to zero and each point a_i of the set E used in defining p-separability, we define a continuous function $f_{ij}(x)$ such that $f_{ij}(x) = 0$ in a closed set A_{ij} containing a_i as an inner point, $f_{ij}(x) = 1$ in a closed set B_{ij} containing every

* "Zum Metrisationsproblem," *Mathematische Annalen,* Vol. 94, pp. 310-311.

x for which $ax \geqq k_j$, and $0 \leqq f_{ij}(x) \leqq 1$ elsewhere in Z. Then the distance between two points x and y is defined by $\sum_{i,j=1}^{i,j=\infty} | f_{ij}(x) - f_{ij}(y) | / 2^{i+j}$.

8. The result of § 6 would appear to be useful in handling upper semi-continuous decompositions of metric spaces into disjoint closed sets. Under certain broad conditions the space whose elements are the closed sets is known to be metric, but the distance between two such elements has no simple relation to the distance between the closed sets in the original space. The following theorem together with § 6 makes it possible to use the distance in the original space as the distance in the new space.

THEOREM. *Let $Z = \{x\}$ be a metric space and $Z = \sum[X]$ be an upper semi-continuous decomposition of Z into disjoint closed sets. Let Z' be the space whose elements are $\{X\}$ and for any two elements X and Y of Z' let XY be the distance between the sets X and Y as measured in Z. Then Z' is a semi-metric space satisfying Axiom IV and, if Z is connected, so is Z'.*

Proof. By the distance between X and Y in Z we mean the lower bound of xy as the point x ranges over X and the point y ranges over Y. If then $XY = 0$, there are sequences $\{x_i\}$ and $\{y_i\}$, chosen from X and Y, respectively, such that $x_i y_i \to 0$. By the definition of upper semi-continuous decompositions it follows that for any $e > 0$ every point of X has a distance from Y less than e and every point of Y has a distance from X less than e. As X and Y are closed, this makes $X = Y$. Hence Z' is semi-metric.

If Axiom IV were not valid, there would be some element A, a constant $k > 0$, and sequences $\{X_i\}$ and $\{Y_i\}$ such that $AX_i \geqq k$, but $AY_i + X_iY_i \to 0$. Take a positive $e < k/3$. Since $AY_i \to 0$ and the decomposition is upper semi-continuous, there is an i' such that every point of Y_i has a distance from A less than e for every $i > i'$. Since $X_iY_i \to 0$, there is an i'' such that for every $i > i''$ there is a point x_i in X_i and a point y_i in Y_i for which $x_i y_i < e$. But then for i greater than both i' and i'' some point of X_i has a distance less than $2e$ from some point of A, and consequently $AX_i < k$, which is false.

If Z' were not connected, it would be the sum of two disjoint non-void sets H' and K' such that $\bar{H}' \cdot K' + H' \cdot \bar{K}' = 0$. In Z let H be the union of the sets $\{X\}$ which are elements of H' and K have a similar relation to K'. If $\bar{H} \cdot K \neq 0$, there would be a point y in K which is the limit of a sequence $\{x_i\}$ of points of H. Now y lies in some element Y and each x_i in some element X_i, whence $X_iY \to 0$. But X_i is an element of H' and Y is an element of K', and $\bar{H}' \cdot K' = 0$, which is a contradiction. Hence $\bar{H} \cdot K = 0$

and, in like manner, $H \cdot \bar{K} = 0$. This is again a contradiction, because Z was connected. Hence Z' is connected.

In connection with this theorem it should be noted that, in general, Axiom V is not satisfied.

9. We shall now investigate certain relationships between semi-metric and topological spaces, defining the latter by these axioms:

A. *Every point x has at least one vicinity $U(x)$ and x lies in $U(x)$.*

B. *If $U(x)$ and $V(x)$ are vicinities of x, there is some $W(x) \subset U(x) \cdot V(x)$.*

C. *For each $U(x)$ there is a $V(x)$ such that, if y lies in $V(x)$, some $U(y) \subset U(x)$.*

4. *If x and y are two distinct points, some $U(x)$ does not contain y.*

Axioms A and B are taken directly from Hausdorff's *Mengenlehre* (pp. 228, 229). Axiom C is the weaker form of Hausdorff's Axiom C suggested by Frechet. This is more convenient, since in semi-metric spaces satisfying Axioms IV or V, a sphere is not necessarily a region, but merely contains a region containing in turn the center. In Axiom 4 "vicinity" has been used instead of "region," since this is more consistent in forming a set of vicinity axioms.

Before proceeding further it will be as well to call attention to two known points which are sometimes overlooked or insufficiently stressed. The first is the fact that it must not be understood that, if some $U(x)$ contains a point y different from x, then $U(x)$ is a vicinity of y. Such an assumption in certain cases vitiates the work. The second is the difference between equivalence and homeomorphism, the former being in some cases an effectively stronger property than the latter. It may be remarked here that the axioms given above are *equivalent* to the corresponding axioms of Hausdorff.

It is clear that relations between semi-metric and topological spaces will involve enumerability axioms of some kind. Consider the following.

9. *There is an equivalent set of vicinities such that every point x has an at most enumerable set of vicinities.*

9'. *If x is a point and $\{U(x)\}$ the set of its vicinities, there is an enumerable sub-set $\{V_i(x)\}$ of these vicinities such that x is the divisor of the set $\{V_i(x)\}$ and each $U(x)$ contains some $V_i(x)$.*

10. *There is an equivalent set of vicinities which is enumerable.*

Axioms 9 and 10 are essentially the same as Hausdorff's axioms with these numbers. It is shown below by an example that Axiom 9′ is effectively weaker than Axiom 9. On the other hand it is readily seen that the restriction of the class of vicinities in Axiom 9′ does not affect the definition of limiting points. This is a case where homeomorphism and equivalence are not the same and it is well illustrated by the following example.

Let Z be the sum of two disjoint sets A and C, where A consists of the points of the interval $0 \leqq x < 1$ and C is a set $\{y\}$ of cardinal number c. For each y let $U(y) = y$; for $x \neq 0$ let $U(x)$ be any open interval of center x contained in A; for $x = 0$ let $U(x)$ be any half-open sub-interval of A whose left end-point is 0 *or* any such sub-interval plus any point of C. Here Axiom 9 is not satisfied. If we restrict the lengths of the sub-intervals to rational numbers and require $U(0)$ to be either such a sub-interval or such a sub-interval plus the point of C corresponding to its length in some correspondence between the rational numbers and an enumerable sub-set of C, Axiom 9 is valid. On the other hand, Axiom 9′ is valid in the original space, since the vicinities of points of C and the vicinities of points of A consisting of sub-intervals of rational length satisfy the requirements. The two spaces are homeomorphic, but not equivalent.

THEOREM I. *Let the topological space Z satisfy Axiom 9′. Then we can take the partial set of vicinities so that for each point $V_1(x) \supseteq V_2(x) \supseteq V_3(x) \supseteq \cdots$.*

Proof. Let $\{U(x)\}$ be the original set and $\{W_i(x)\}$ be any partial set satisfying the requirements of Axiom 9′. Take $V_1(x) = W_1(x)$. Now $W_1(x) \cdot W_2(x) \supset$ some $U(x) \supset$ some $W_{i_2}(x)$. Set $V_2(x) = W_{i_2}(x)$. Likewise the divisor of the first $i_2 + 1$ vicinities $\{W_i(x)\}$ contains some $U(x)$, which in turn contains some $W_{i_3}(x)$; this we take for $V_3(x)$. Continue this process indefinitely. Clearly every $U(x)$ contains some $V_n(x)$ and x is the divisor of the monotone descending sequence $\{V_n(x)\}$.

THEOREM II. *Let the topological space Z satisfy Axiom 9. Then there is an equivalent set of vicinities so that for each point $V_1(x) \supseteq V_2(x) \supseteq \cdots$.*

This is proved in the same way as Theorem I.

10. It is apparent that there is an intimate connection between Axioms III, IV, and V of semi-metric spaces and Tietze's separation axioms * for topological spaces. But in studying this connection we meet the following

* See Hausdorff, *Mengenlehre*, p. 229, Axioms 5-8.

difficulty. If in a semi-metric space the sequence of points $\{a_i\}$ converges to a point a, there is for every $r > 0$ an i' such that a sphere of center a and radius r contains every a_i for which $i > i'$ and for such values of i the spheres of radius r and centers a_i all contain a. But this does not hold for topological spaces. There, although each vicinity of a contains every a_i for i greater than some i', it may be that *no* vicinity of any a_i contains a. Therefore it seems well to the author to propose the following axioms in place of Hausdorff's Axioms 5-8 for use in topological spaces satisfying Axioms 9 or 9′:

5′. *For every pair of points a and b and every integer n there is an integer $m = g(a, b, n)$ such that m increases indefinitely with n and the relation $V_n(a) \cdot V_n(b) \neq 0$ implies that b lies in $V_m(a)$ and a lies in $V_m(b)$.*

6′. *For each point a and each integer n there is an integer $m = g(a, n)$ such that m increases indefinitely with n and the relation $V_n(a) \cdot V_n(b) \neq 0$ implies that b lies in $V_m(a)$ and a lies in $V_m(b)$.*

7′. *For each integer n there is an integer $m = g(n)$ such that m increases indefinitely with n and the relation $V_n(a) \cdot V_n(b) \neq 0$ implies that b lies in $V_m(a)$ and a lies in $V_m(b)$.*

It is a simple matter to show that Axiom 5′ implies Hausdorff's Axiom 5. For Axiom 6′ we get the following theorem, which is analogous to a theorem of Tychonoff.*

THEOREM I. *Let Z be a topological space satisfying Axioms A, B, C, 4, and 9 or 9′. If it also satisfies Axiom 6′, it satisfies Hausdorff's Axioms 6, 7, and 8.*

Proof. Let our vicinities be monotone descending as in the theorems of § 9. Let A and B be two point-sets such that $\bar{A} \cdot B + A \cdot \bar{B} = 0$. Let a be a fixed point not in the closed set $\bar{B}$ and b be any point of $\bar{B}$. Then there is an n such that $V_n(a) \cdot V_n(b) = 0$. If n is unbounded as b ranges over $\bar{B}$, there is for each n a point b_n in $\bar{B}$ such that $V_n(a) \cdot V_n(b_n) \neq 0$. If $m = g(a, n)$ as in Axiom 6′, $V_m(a)$ contains b_n. But m increases indefinitely with n; hence every $V_m(a)$ contains points of $\bar{B}$, a contradiction. Therefore for some n we have $V_n(a) \cdot V_n(b) = 0$ for every point b in $\bar{B}$.

In consequence of this result there is for each integer i a sub-set A_i of A such that for each point a in A_i and each point b in $\bar{B}$, $V_i(a) \cdot V_i(b) = 0$. The set A_i may be void for a particular value of i, but $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ and A is the union of the sets $\{A_i\}$. Likewise, B is the union of a monotone

* "Über einen Metrisationssatz" von P. Urysohn, *Mathematische Annalen*, Vol. 95, pp. 139-141.

increasing sequence of sets $\{B_i\}$, such that for each b in B_i and each a in $\bar{A}$, $V_i(a) \cdot V_i(b) = 0$.

Each $V_i(a)$ contains a region $v_i(a)$ which contains a, by Axiom C, and likewise for $V_i(b)$. Let U_i be the union of the regions $\{v_i(a)\}$ as a ranges over A_i, and W_i be the union of the regions $\{v_i(b)\}$ as b ranges over B_i. Let i and j be any two integers and $i \leqq j$. If a lies in A_i, $V_i(a) \cdot V_i(b) = 0$ for every b in $\bar{B}$, and consequently in B_j. Since $V_j(b) \subseteq V_i(b)$, we have $V_i(a) \cdot V_j(b) = 0$ for every a in A_i and b in B_j. Consequently $U_i \cdot W_j = 0$ for this case, and similar reasoning establishes the same fact for the case that $i > j$. If then R and S are the unions of the sets $\{U_i\}$ and $\{W_i\}$, respectively, we have $A \subset R$, $B \subset S$, and $R \cdot S = 0$. As R and S are obviously regions, we have shown that A and B are contained in disjoint regions, which is the requirement in Hausdorff's Axiom 8. A fortiori, Axioms 6 and 7 are also valid.

THEOREM II. *Let Z be a topological space satisfying Axioms A, B, C, 4, 6′, and 9 or 9′, and $\{a_i\}$ a sequence of points converging to a. Then for each m there is an i_m such that a_i lies in $V_m(a)$ and a lies in $V_m(a_i)$ for every $i > i_m$.*

Proof. As usual we assume that our vicinities are monotone descending in accordance with the theorems of § 9. Since by Axiom 6′, $m = g(a, n)$ increases indefinitely with n, there is for each integer m an integer n such that $V_n(a) \cdot V_n(b) \neq 0$ implies that a lies in $V_m(b)$ and b lies in $V_m(a)$.

Since $a_i \rightarrow a$, there is an i_n such that each a_i lies in $V_n(a)$ for $i > i_n$. But then $V_n(a) \cdot V_n(a_i) \neq 0$. Consequently the previous paragraph is applicable and we have the theorem on writing i_n as i_m.

11. THEOREM. *Let Z be a semi-metric space satisfying Axiom IV(V). Let $r_1 > r_2 > r_3 \cdots$ and $r_i \rightarrow 0$. For each point a of Z let $U_i(a)$ denote a sphere of center a and radius r_i. If these spheres are taken as vicinities, Z is a topological space satisfying Axioms A, B, C, 4, and 9, and also 6′(7′).*

Proof. It is clear that Axioms A, B, 4, and 9 are satisfied. Now take a fixed $U_i(a)$. Then for some $j > i$, it follows from § 4 that $U_j(a)$ contains only inner points of $U_i(a)$ and so for any point x in $U_j(a)$ some $U_k(x) \subset U_i(a)$. Hence Axiom C is also valid.

Now suppose that Axiom 6′ were not valid. Then there would be a fixed integer m and a sequence of points $\{b_n\}$ such that $U_n(a) \cdot U_n(b_n) \neq 0$, but for every n either a would not lie in $U_m(b_n)$ or b_n would not lie in $U_m(a)$. The former statement requires the existence of a sequence $\{c_n\}$ such that

$ac_n \to 0$ and $b_nc_n \to 0$. But then by Axiom IV we have $ab_n \to 0$. Then for some n_0 and every $n > n_0$, $ab_n < r_m$, which contradicts the second statement. Hence Axiom 6′ is valid.

Suppose now that Axiom V is valid, but Axiom 7′ is not. Then there is a fixed m and sequences $\{a_n\}$ and $\{b_n\}$ such that $U_n(a_n) \cdot U_n(b_n) \neq 0$, but either b_n is not in $U_m(a_n)$ or a_n is not in $U_m(b_n)$. Then for a sequence $\{c_n\}$, we have $a_nc_n + b_nc_n \to 0$, whence $a_nb_n \to 0$. This gives a contradiction as above.

Remark. The above theorem is a fortiori true if 9′ is substituted for 9.

12. THEOREM. *Let Z be a topological space satisfying Axioms A, B, C, 4, and 9. If also Axiom 6′(7′) is satisfied, Z is equivalent to a semi-metric space satisfying Axiom IV(V).*

Proof. In accordance with § 9 we assume that the vicinities of each point form a monotone descending sequence of sets. For a pair of points a and b set $f_n(a, b) = 0$ if b lies in $V_n(a)$ and $f_n(a, b) = 1$ if b is not in $V_n(a)$. Likewise define $f_n(b, a)$. Let $d_n(a, b) = d_n(b, a) = [f_n(a, b) + f_n(b, a)]/2^n$ and $ab = ba = \sum_1^\infty d_n(a, b)$. Let Z' be a space having the same points as Z and distances defined in this manner.

It follows at once from Axiom 6′ that there is an integer n' such that $V_n(a) \cdot V_n(b) = 0$ if $n \geqq n'$. Hence $ab \geqq \sum_{n'}^\infty 2/2^n = 1/2^{n'-2} > 0$ and so Z' is semi-metric.

Now let a be any point and $r > 0$. Take m so that $r > 1/2^{m-1}$. In consequence of Axiom 6′ there is an n such that, if b lies in $V_n(a)$, then b lies in $V_m(a)$ and a lies in $V_m(b)$. Hence $ab < \sum_{m+1}^\infty 2/2^i = 1/2^{m-1} < r$. That is, each sphere of radius r contains some $V_n(a)$. On the other hand, let $V_m(a)$ be any vicinity. If $r < 1/2^m$ and $ab < r$, $d_n(a, b) = 0$ for every $n \leqq m$. Hence by the definition of $d_n(a, b)$, b lies in $V_n(a)$ and a in $V_n(b)$ for every $n \leqq m$. Thus every vicinity of a contains some sphere of center a. Hence we have proved that the spaces are *equivalent*.

If Axiom 6′ does not imply Axiom IV, there are two sequences $\{b_i\}$ and $\{c_i\}$ and a constant $k > 0$ such that $ab_i \geqq k$ and $ac_i + b_ic_i \to 0$. These relations show that for a fixed n' there is an i' such that $d_n(a, c_i) = 0$ and $d_n(b_i, c_i) = 0$ for every $i \geqq i'$ and every $n \leqq n'$. Then for such values $V_n(c_i)$ contains a and b_i, and c_i lies in both $V_n(a)$ and $V_n(b_i)$. If $m = g(a, n)$

as defined in Axiom 6′, this means that b_i lies in $V_m(a)$ for $i \geqq i'$. As m increases indefinitely with n, we have b_i approaching a, which is impossible.

If Axiom 7′ does not imply Axiom V, we have three sequences $\{a_i\}$, $\{b_i\}$, and $\{c_i\}$, and a $k > 0$, such that $a_ib_i \geqq k$ and $a_ic_i + b_ic_i \to 0$. As above, we have for each n' an i' such that $V_n(a_i) \cdot V_n(b_i) \neq 0$ for $i \geqq i'$ and $n \leqq n'$. Let $m = g(n)$ as defined in Axiom 7′. Then $d_n(a_i, b_i) = 0$ for $n \leqq m$; and, as m increases indefinitely with n, this means that $a_ib_i \to 0$, a contradiction.

Corollary. *Let Z be a topological space satisfying Axioms A, B, C, 4, and 9′. If also Axiom 6′(7′) is satisfied, Z is homeomorphic with a semi-metric space satisfying Axiom IV(V).*

For by § 9 the space Z is homeomorphic with a topological space satisfying Axioms A, B, C, 4, and 9.

Remark. A study of the proof of the theorem of § 6 shows that in the above theorem Axiom 6′ is sufficient for equivalence with a semi-metric space satisfying Axiom V, and an analogous remark is true for the corollary.

13. In the theorem and corollary of the previous section a distinction has been drawn between homeomorphism and equivalence. The same thing is necessary in connection with Urysohn's theorem regarding the equivalence of a topological space satisfying Axioms A, B, C, 6, and 10 to a metric space, which has been referred to in § 7. In his proof it is tacitly assumed that every vicinity containing a point is a vicinity of that point; without that assumption the proof given does not apply and in fact there is homeomorphism and not equivalence. This in part explains the necessity for such axioms as 6′ and 7′.

Since it has been shown that a semi-metric space satisfying Axiom IV is homeomorphic with a metric space, it follows from the previous section that this is also true for a topological space satisfying Axioms A, B, C, 4, 6′, and 9′.

CONCERNING HEREDITARILY LOCALLY CONNECTED CONTINUA.

By G. T. Whyburn.

A continuum every subcontinuum of which is locally connected is said to be *hereditarily locally connected.* The principal contribution of the present paper is the establishing of the proposition that *Every hereditarily locally connected, compact and metric continuum is a rational curve in the Menger-Urysohn* * *sense,* that is, each point of such a continuum M is contained in arbitrarily small neighborhoods having countable boundaries relative to M. This theorem was proved formerly † by the present author for subcontinua of the plane, but the demonstration in the present article is independent of the containing space and is therefore valid for subcontinua of any compact metric space. It is thus demonstrated that in the Menger–Urysohn classification of curves, the hereditarily locally connected continua form a distinct class occupying an intermediate position between the class of all regular curves (= continua each of whose points is contained in arbitrarily small neighborhoods with finite boundaries) and the class of all rational curves. In other words, all regular curves are hereditarily locally connected, but not conversely; and all hereditarily locally connected continua are rational curves, but not conversely.

In the course of the demonstration of the proposition announced above, the author has found and used a number of strong properties of hereditarily locally connected continua, each of which, incidentally, characterizes these continua among the compact metric continua. Proofs for these properties, together with some of their corollaries, form § 2 of the present paper. In § 3 there is given two lemmas of a general character which also are needed in the proof of the main theorem of the paper, given in § 4.

We shall employ the usual terminology and notation of the theory of sets. Our hypotheses ordinarily concern a compact metric continuum M and its subsets, and in such cases we shall consider M as a space and shall speak of the open subsets of M as neighborhoods or open sets. If V is such a set, $F(V)$ will denote the boundary of V, i. e., the point set $\bar{V} - V$. A component of a set K is a connected subset of K which is contained in no other

* See Menger, *Mathematische Annalen,* Vol. 95 (1925), pp. 272-306 and Urysohn, *Verhandelingen der Akademie te Amsterdam,* Vol. 13 (1927), No. 4.

† See *Bulletin of the American Mathematical Society,* Vol. 36 (1930), pp. 522-524.

connected subset of K. The quasi-component of a set K containing the point p of K consists of p together with all points x such that K is not the sum of two mutually separated sets, one containing p and the other x. A collection of sets will be called a *null family* provided that all save a finite number of these sets are of diameter less than any preassigned positive number. A countable sequence of sets whose elements form a null family will be called a *null sequence.*

2. *Properties equivalent to hereditary local connectivity.*

(1). *In order that the compact metric continuum M should be hereditarily locally connected it is necessary and sufficient that if K is any subset of M and G is any collection of open sets covering some subset H of K, the boundary of no one of which contains a point of K, then there exists a null sequence $V_1, V_2, V_3, \cdots$ of mutually exclusive open sets covering H each of which is a subset of some element of G and its boundary is a subset of the sum of the boundaries of a finite number of the sets of G.*

The condition is necessary. For by the Lindelöf Theorem there exists a countable sequence $G_1, G_2, G_3, \cdots$ of the sets of G whose sum covers H. Set

$$G_1 = U_1,\ G_2 - G_2 \cdot \bar{G}_1 = U_2,\ G_3 - G_3 \cdot (\bar{G}_1 + \bar{G}_2) = U_3, \cdots,\ G_n - G_n \cdot \sum_1^{n-1} \bar{G}_i = U_n, \cdots.$$

Then clearly $U_1, U_2, \cdots$ is a sequence of mutually exclusive open sets covering H, and for each n,

$$\text{(i)}\quad F(U_n) = F[G_n - G_n \cdot \sum_1^{n-1} \bar{G}_i] \subset F(G_n) + \sum_1^{n-1} F(G_i) = \sum_1^{n} F(G_i).$$

Now set $U = \sum_1^{\infty} U_n$. Then U is an open subset of M, and hence the components of U may be arranged into a sequence $V_1, V_2, V_3, \cdots$, which clearly must be a null sequence, since M is hereditarily locally connected. Since the sets $[U_n]$ are mutually exclusive, it follows that for each n there exists an i such that $V_n \subset U_i$ and $F(V_n) \subset F(U_i)$. Hence $V_n \subset G_i$, and by (i) we have $F(V_n) \subset \sum_1^{i} F(G_j)$, which completes the proof.

That the condition is sufficient follows at once from the fact * that every continuum M which is not hereditarily locally connected contains an infinite sequence $N_1, N_2, N_3, \cdots$ of mutually exclusive continua, all of diameter greater

* See R. L. Moore, *Bulletin of the American Mathematical Society*, Vol. 29 (1923), p. 296; also C. Zarankiewicz, *Fundamenta Mathematicae*, Vol. 9, p. 134.

than some $\epsilon > 0$, which converges sequentially to a limit continuum N having no point in common with the continua $[N_i]$. For, taking $K = H = \sum_1^\infty N_i$ and, for each i, letting G_i denote the set of all points x in M such that $\rho(x, N_i) < (1/3)\rho(N_i, H - N_i)$, then the collection $[G_i]$ covers H and the boundary of no one of these sets contains a point of K, but clearly no null sequence exists satisfying the terms of our condition.

(2). *In order that the compact metric continuum M be hereditarily locally connected it is necessary and sufficient that if K is any subset of M, and p is any point of a quasi-component C of K, and R is any neighborhood of p, then there exist a neighborhood U of p such that $R \cdot C \subset U \subset R$ and $F(U) \cdot K \subset F(R) \cdot C$.*

The condition is necessary. For let H denote the set $(K - C) \cdot F(R)$. Then since no point of H belongs to C, there exists, for each point x of H, a separation of K into two mutually separated sets one containing x and the other C; and hence there exists an open set G_x containing x but not C and such that $F(G_x) \cdot K = 0$. The collection G of all sets $[G_x]$ for all points x of H covers H, and the boundary of no one of these sets contains a point of K. Therefore, by (1), there exists a null sequence $V_1, V_2, V_3, \cdots$ of open sets covering H each of which contains at least one point of H and is a subset of some G_x and has no boundary point in K. Therefore $\sum_1^\infty V_i$ contains H but contains no point of C, and $K \cdot \sum_1^\infty F(V_i) = 0$. But since $[V_i]$ is a null sequence and since each V_i contains at least one point of $F(R)$, it follows that $F(\sum V_i) \subset \sum F(V_i) + F(R)$, and hence $K \cdot F(\sum V_i) \subset K \cdot F(R)$, because $K \cdot \sum F(V_i) = 0$. Then since H, which is $= F(R) \cdot (K - C)$, is a subset of $\sum V_i$, therefore $F(\sum V_i) \cdot K \subset F(R) \cdot C$.

Now set $V = \sum_1^\infty V_i$. We have just shown that V contains H but no point of C and that $F(V) \cdot K \subset F(R) \cdot C$. Set $U = R \cdot (M - \bar{V})$. Then $R \cdot C \subset U \subset R$ and $F(U) \subset F(V) + [F(R) - H]$, and therefore

$$F(U) \cdot K \subset F(V) \cdot K + F(R) \cdot C \subset F(R) \cdot C.$$

To prove the sufficiency of the condition, take the sets $N, N_1, N_2, \cdots$ as in the proof of the sufficiency part of (1). Let $K = N + \sum_1^\infty N_i$, let p be a point of N, and let R be a neighborhood of p of diameter $< \epsilon/2$. Then although N is a quasi-component of K, there can exist no neighborhood U of

p satisfying the terms of our condition, because any such U would contain a point of some N_i and since $\delta(N_i) > \epsilon$, we would have $N_i \cdot F(U) \neq 0$.

(3). *In order that the compact metric continuum M should be hereditarily locally connected it is necessary and sufficient that the quasi-components and the components of any subset of M be identical.*

The condition is necessary. For let K be any subset of a hereditarily locally connected continuum M. Clearly it is sufficient to prove that every quasi-component of K is connected. Suppose, on the contrary, that some quasi-component C of K is the sum of two mutually separated sets C_1 and C_2. Then there exists an open set R containing C_1 but no point of C_2 and such that $F(R) \cdot C = 0$. Now by (2) there exists an open set U such that $R \cdot C \subset U \subset R$ and $F(U) \cdot K \subset F(R) \cdot C = 0$. Thus U contains C_1 but not C_2, which is impossible since C is a quasi-component of K and $F(U) \cdot K = 0$.

The condition is also sufficient. For consider the sets $N, N_1, N_2, \cdots$ as in the preceding proofs. Let a and b be distinct points of N, and let $K = a + b + \sum_1^\infty N_i$. Then obviously a is a component of K, whereas $a + b$ is the quasi-component of K containing a. Thus our condition is contradicted in any continuum which is not hereditarily locally connected.

(4). *In order that the compact metric continuum M should be hereditarily locally connected it is necessary and sufficient that the components of any subset of M form a null family.*

The condition is necessary. For let M be any hereditarily locally connected continuum and suppose, contrary to our theorem, that there exists an infinite sequence $K_1, K_2, K_3, \cdots$ of distinct components of some subset of M all of which are of diameter greater than some given positive number ϵ. Set $K = \sum_1^\infty K_n$. Then for each n, K_n is a component of K. By (3), K_n is also a quasi-component of K, for each n. Thus there exists a separation of K into two mutually separated sets H_1 and H_2 containing K_1 and K_2 respectively, and hence there exist two mutually exclusive open sets G_1 and Q_1 containing H_1 and H_2 respectively and such that the boundary of neither of these sets contains a point of K. One of these sets, say Q_1, contains infinitely many of the sets $[K_n]$. Likewise there exists a separation of $K \cdot Q_1$ into two non-vacuous mutually separated sets L_1 and L_2, and hence there exist two mutually exclusive open subsets G_2 and Q_2 of Q_1 containing L_1 and L_2 respectively, the boundary of neither of which contains a point of K. One of these sets, say Q_2,

contains infinitely many of the sets $[K_n]$. There exists a separation of $K \cdot Q_2$, and so on. Continuing this process indefinitely, we obtain an infinite sequence of mutually exclusive open sets $G_1, G_2, G_3, \cdots$, such that for each i, G_i contains some component K_{n_i} of K and $F(G_i) \cdot K = 0$. Set $H = \sum_1^\infty K_{n_i}$. Then, by (1), there exists a null sequence $V_1, V_2, \cdots$ of open sets covering H each of which is a subset of some G_i and is such that its boundary contains no point of K. But since the sets G_i are mutually exclusive, no set V_n can contain more than one of the sets K_{n_i}. But then clearly the fact that $\delta(K_{n_i}) > \epsilon$ for all i's contradicts the fact that $V_1, V_2, V_3, \cdots$ is a null sequence.

The sufficiency of the condition is an immediate consequence of the existence of the sets $N, N_1, N_2, N_3, \cdots$, as previously defined, in any continuum which is not hereditarily locally connected, because for each i, N_i is a component of $\sum_1^\infty N_n$ and $\delta(N_i) > \epsilon$, which contradicts our condition.

(5). *In order that the compact metric continuum M should be hereditarily locally connected it is necessary and sufficient that every connected subset of M should be locally connected.* (Theorem of R. L. Wilder.)*

The sufficiency of the condition is obvious. To prove the necessity, let us suppose, on the contrary, that some connected subset H of the hereditarily locally connected continuum M is not locally connected at one of its points p. Then there exists a neighborhood E of p and an infinite sequence $p_1, p_2, p_3, \cdots$ of distinct points of $H \cdot E$ such that $\rho(p_i, p) < (1/2)\rho[p, F(E)]$ for every i, and no two points of this sequence lie in the same component of $K = H \cdot \bar{E}$. For each i let C_i denote the component of K containing p_i. Since, by (4), the components of K form a null family, it follows that for some i, $C_i \cdot F(E) = 0$. Then, applying (2), we obtain a neighborhood U of p_i such that $E \cdot C_i \subset U \subset E$ and $F(U) \cdot K \subset F(E) \cdot C_i = 0$. Since $F(U) \subset \bar{E}$, therefore we have $0 = F(U) \cdot K = F(U) \cdot H \cdot \bar{E} = F(U) \cdot H$. But then H is the sum of the two mutually separated sets $H \cdot U$ and $H - H \cdot U$, contrary to the fact that H is connected.

COROLLARIES. *Let M be any compact, metric, and hereditarily locally connected continuum and let K be any subset of M. Then:*

* See R. L. Wilder, *Proceedings of the National Academy of Sciences*, Vol. 15 (1929), p. 616. This theorem and its proof are included in the present paper for the sake of completeness of the treatment. The proof given is the author's own.

(a) *If C is any component of K and R is any open set, there exists an open set U such that $R \cdot C \subset U \subset R$ and $F(U) \cdot K \subset F(R) \cdot C$.* [Consequence of (2) and (3)].

(b) *If K is of dimension > 0 at any one of its points p, then p lies in a non-degenerate connected subset of K.* [Consequence of (2) and (3)].

(c) *Either K is of dimension zero, or it contains non-degenerate connected sets but no continua, or it contains continua.* [Consequence of (b)].

Remark. Corollary (c) states that the subsets of any hereditarily locally connected continuum fall into three mutually exclusive classes as follows: (α) sets containing continua, (β) punctiform sets which contain non-degenerate connected sets, and (γ) zero-dimensional sets. Sets of all three classes are known * to exist in hereditarily locally connected continua, and it is now definitely established that there are no others. This classification is actually a restrictive one, because it tells us, for example, that sets such as the totally disconnected one-dimensional sets are not to be found among the subsets of hereditarily locally connected continua. The classification is all the more restrictive in view of the theorem of Wilder [see (5) above] that every connected subset of such a continuum is locally connected.

3. LEMMA 1. *In any separable metric space R there exists a countable set of points D such that if p and q are any two points whatever which can be separated by some countable set, then p and q can be separated by some subset of D.*†

Proof. There exists a countable set of points $Q = x_1 + x_2 + x_3 + \cdots$ which is dense in R. Order all possible pairs of points x_i, x_j of Q such that x_i and x_j can be separated by some countable set into a sequence $P_1, P_2, P_3, \cdots$. For each n there exists a countable set of points E, which we may suppose closed,‡ which separates the two points x_i and x_j in R. There exists a positive real number a such that E also separates in R the point sets $V_a(x_i)$ and

* See similar remarks by the author concerning the subsets of regular curves in *Monatshefte für Mathematik und Physik*, Vol. 38 (1931), and note references therein to examples by Knaster, Kuratowski, Sierpinski, and Mazurkiewicz.

† It is evident from the proof of this lemma that the same argument suffices to establish the following general theorem: *If S is any class of closed subsets of a separable metric space R, there exists a countable sub-class $[S_i]$ of S such that each pair of points which may be separated by some set of the class S may also be separated by some set of the class $[S_i]$.* This general proposition together with some of its consequences will be considered by the author in a later paper.

‡ See Tietze, *Mathematische Annalen*, Vol. 88, p. 310.

$V_a(x_j)$, where $V_r(x)$ denotes in general the set of all points of the space whose distances from the point x are less than the positive real number r. With the aid of the Dedekind Cut-Postulate it is seen that there exists a number a_n, $0 < a_n \leq (1/2)\rho(x_i, x_j)$, such that for every positive number $a < a_n$, but for no number $> a_n$, there exists some countable set which separates the point sets $V_a(x_i)$ and $V_a(x_j)$. For each n let D_n denote some countable set which separates the point sets $V_{b_n}(x_i)$ and $V_{b_n}(x_j)$, where $b_n = a_n - 1/n$. Let $D = \sum_1^\infty D_n$. Then D has the required properties.

For let p and q be any two points which can be separated in R by some countable set E. We may suppose E closed, and hence there exists a positive number u such that E also separates the sets $V_u(p)$ and $V_u(q)$. Since Q is dense in R, it follows that there exists an integer $n > 8/u$ such that the points x_i and x_j of the pair P_n satisfy the relations $\rho(x_i, p) < u/8$ and $\rho(x_j, q) < u/8$. Hence $V_{u/2}(x_i) \subset V_u(p)$ and $V_{u/2}(x_j) \subset V_u(q)$, and therefore $a_n \geq u/2$. Thus $b_n = a_n - 1/n > a_n - u/8 > u/2 - u/8 > u/4$, because $1/n < u/8$. Hence $V_{b_n}(x_i) \supset p$ and $V_{b_n}(x_j) \supset q$, and therefore the set D_n, which is a subset of D, separates p and q in R.

Definitions. Any connected open subset of a locally connected space N will be called a *region* in that space. A region R is said to *join* two point sets A and B provided R contains at least one point of A and at least one point of B. Two regions R_1 and R_2 will be said to be *strongly separated* provided that they have no points and no boundary points in common, i. e., $\bar{R}_1 \cdot \bar{R}_2 = 0$.

LEMMA 2. *If N is any connected and locally connected metric space which has no cut point and A and B are any two mutually exclusive non-degenerate subsets of N, then there exist two strongly separated regions in N joining A and B.*

Proof. We may suppose that A and B are closed, for obviously they contain closed and non-degenerate subsets. Let p denote some point of A. Since $N - p$ is connected, there exists * a region R_{ab} joining A and B and such that $p \cdot \bar{R}_{ab} = 0$. There exists a region R_{ax} containing p and such that $\bar{R}_{ax} \cdot \bar{R}_{ab} = 0$. Thus there exist points x of N such that two strongly sepa-

* See R. L. Wilder, *Bulletin of the American Mathematical Society*, Vol. 34 (1928), pp. 649-655. It is only necessary to cover $N - p$ with a set of regions no one of which contains p or has p on its boundary and then take R_{ab} equal to the sum of the elements of a finite simple chain of these regions joining some point a of A and some point b of B.

rated regions R_{ab} and R_{ax} exist joining A and B and A and x respectively. Let S denote the set of all such points x of N. We shall show that $S = N$. Suppose this is not so. Then since obviously S is open in N and N is connected, it follows that some point y of $N - S$ is a limit point of S. Since $N - y$ is connected, it follows just as above in the case of p that there exist two strongly separated regions G_{ab} and G such that G_{ab} joins A and B and G contains y. The region G contains a point x of S; and hence there exist two strongly separated regions R_{ab} and R_{ax} joining A and B and A and x respectively.

Now let R_1 and R_2 respectively denote components of $R_{ab} - \bar{G} \cdot R_{ab}$ and $R_{ax} - \bar{G} \cdot R_{ax}$ each of which contains at least one point of A. Inasmuch as N is locally connected and y does not belong to S, it follows that $\bar{G} \cdot \bar{R}_1 \neq 0 \neq \bar{G} \cdot \bar{R}_2$. Let K denote the set $A + \bar{R}_1 + \bar{R}_2$, and let U be a component of $G_{ab} - K \cdot G_{ab}$ which contains at least one point of B. At least one point f of K is a limit point of U, because N is locally connected. Let Z denote one of the sets R_1 and R_2 such that $\bar{Z}$ does not contain f and let T denote the other one of these sets. Let g denote some point of $\bar{G} \cdot \bar{Z}$. Since f cannot belong to $\bar{G}$, there exist strongly separated regions U_f and U_g containing f and g respectively and such that $\bar{U}_f \cdot (\bar{G} + \bar{Z}) = 0$ and $\bar{U}_g \cdot (\bar{T} + \bar{U}) = 0$. Clearly $U + U_f + T$ contains a region V joining A and B and $Z + U_g + G$ is a region V_{ay} joining A and y, and the regions V and V_{ay} have no point in common. But V contains a region V_{ab} joining A and B and such that $\bar{V}_{ab} \subset V$; for it is only necessary to cover V with regions $[V_p]$ each lying together with its boundary wholly in V, and then take V_{ab} equal to the sum of the elements of a finite simple chain of these regions $[V_p]$ joining some point a of A and some point b of B. This is impossible, because the regions V_{ab} and V_{ay} are strongly separated and join A and B and A and y respectively, contrary to the fact that y does not belong to S. Therefore $S = M$. Accordingly S contains a point x of B, and thus there exist two strongly separated regions R_{ab} and R_{ax} joining A and B. Q. E. D.

4. THEOREM. *Every hereditarily locally connected, compact and metric continuum is a rational curve.*

Proof. Suppose, on the contrary, that some continuum M exists satisfying our hypothesis but which is not rational. By Lemma 1 there exists a countable subset D of M such that if any two points p and q can be separated in M by some countable set, then p and q can be separated by some subset of D. Now there exists at least one non-degenerate component C of $M - D$. For if every point of $M - D$ is a component of $M - D$, then by § 2, results

(2) and (3) or Corollary (b), it follows that $M - D$ is zero-dimensional at every point and hence that M is rational at every point of $M - D$, for D is countable. But D being countable, it follows that M is rational at all of its points, contrary to hypothesis. Thus there exists a non-degenerate component C of $M - D$.

Now no two points of $\bar{C}$ can be separated in $\bar{C}$ by any countable set of points. For if, on the contrary, some two points p' and q' of $\bar{C}$ can be separated in $\bar{C}$ by some countable set E, it follows that some two points p and q of C can be separated in $\bar{C}$ by E; and hence there exists a neighborhood R of p such that $\bar{R}$ does not contain q and $F(R) \cdot C$ is countable. By § 2, result (2), there exists a neighborhood U of p such that $U \subset R$ and $F(U) \cdot (M - D) \subset F(R) \cdot C$, and hence such that $F(U) \cdot (M - D)$ is countable. But $F(U) \cdot M = F(U) \cdot (M - D) + F(U) \cdot D$, and hence $F(U) \cdot M$ is countable. Thus p and q are separated in M by the countable set $F(U) \cdot M$, and therefore they can be separated in M by some subset of D. But this is impossible, because C is connected and contains both p and q but contains no point of D. Consequently no two points of $\bar{C}$ can be separated in $\bar{C}$ by any countable set.

For convenience of notation we set $\bar{C} = N$. Then N is a hereditarily locally connected continuum no two points of which can be separated in N by any countable set of points. Now the local separating points * of any connected subset H of M which are not rational points of H must be countable. For if G is any uncountable set of local separating points of H, then since, by (5) in § 2, H is locally connected, it follows that there exists a region R in H and an uncountable subset E of G every point of which is a cut point of R. Now it is a consequence of a theorem of the author's † that there exists a point p of E and a countable subset D of R such that p is a component of $R - D$. Then by § 2, Corollary (b), $R - D$ is zero-dimensional at p, and therefore both R and H are rational at p. Thus every such set G contains a point in which H is rational, and accordingly the local separating points of H which are not rational points of H are countable. Now since N is not rational in any of its points, it follows that the set D_0 of all local separating points of $N = N_0$ is countable. Hence $N_1 = N - D_0$ is connected, because D_0 cannot separate any two points of N. Now N_1 cannot be rational at any one of its points, because a point of rationality of N_1 would be also a point of rationality of N, since D_0 is countable. Thus it follows that the set

* The point p of a connected and locally connected set H is a local separating point of H provided that p is a cut point of some region in H.

† See my paper "Non-Separated Cuttings of Connected Sets," *Transactions of the American Mathematical Society*, Vol. 33 (1931).

D_1 of all local separating points of N_1 is countable. Hence $N_2 = N_1 - D_1$ is connected, and so on. For any ordinal number a of the first or second class, let us suppose we have defined the sets N_β and D_β for all ordinal numbers $\beta < a$. Then $\sum_{\beta<a} D_\beta$ is countable, and therefore $N - \sum_{\beta<a} D_\beta$ is connected. Set $N - \sum_{\beta<a} D_\beta = N_a$ and let D_a denote the set of all local separating points of N_a. It follows just as above that D_a is countable. Thus we have defined the sets N_a and D_a for all ordinal numbers a of the first and second classes.

We shall now show that for some a, $D_a = 0$. Suppose, on the contrary, that $D_a \neq 0$ for every a of the first and second class. Then for each a there exists a point p_a of D_a, and p_a is a local separating point of N_a but not of N_β for any $\beta < a$. Since the set of points $[p_a]$ is uncountable and since for each a there exists an $\epsilon_a > 0$ such that p_a is a cut point of the component of $N_a \cdot V_{\epsilon a}(p)$ containing p_a, it follows that there exists some $\epsilon > 0$, a point p, and an uncountable subset $[p_{a_a}]$ of $[p_a]$, where $a_\beta < a_a$ if $\beta < a$, such that for each a_a, $\rho(p, p_{a_a}) < \epsilon/4$ and p_{a_a} is a cut point of the component C_a of $N_{a_a} \cdot V_\epsilon(p)$ which contains p_{a_a}. Now inasmuch as $C_1 - p_{a_1}$ has at least two distinct components, there exists at least one component F_1 of $C_1 - p_{a_1}$ and an infinite subset E_1 of $[p_{a_a}]$ such that $E_1 \cdot F_1 = 0$. Let $p_{a_{n_2}}$ be the first point in the sequence $[p_{a_a}]$ following p_{a_1} which belongs to E_1. Then, just as before, there exists at least one component F_2 of $C_{n_2} - p_{a_{n_2}}$ and an infinite subset E_2 of E_1 such that $E_2 \cdot F_2 = 0$. Let $p_{a_{n_3}}$ be the first point in the sequence $[p_{a_a}]$ following $p_{a_{n_2}}$ which belongs to E_2. There exists a component F_3 of $C_{n_3} - p_{a_{n_3}}$ and an infinite subset E_3 of E_2 such that $E_3 \cdot F_3 = 0$, and so on. Continuing this process indefinitely we obtain an infinite sequence of sets $F_1, F_2, F_3, \cdots$. Now since for each i, F_i is a component of $C_{n_i} - p_{a_{n_i}}$, since $\rho(p, p_{a_{n_i}}) < \epsilon/4$, and since C_{n_i} is connected and locally connected and clearly $\bar{C}_{n_i} \cdot F[V_\epsilon(p)] \neq 0$, it follows at once that $\delta(F_i) > \epsilon/2$ for all i's. Now for each $j > i$, we have $N_{a_{n_j}} \subset N_{a_{n_i}} - p_{a_{n_i}}$; and therefore $C_{n_j} \subset C_{n_i} - p_{a_{n_i}}$. Thus since $p_{a_{n_j}} \subset E_j$ and $E_j \cdot F_i = 0$, it follows that $C_{n_j} \cdot F_i = 0$; and as $F_j \subset C_{n_j}$, therefore $F_i \cdot F_j = 0$ for every pair of integers i and j. Since for each i it is true that for every $j > i$, $C_{n_j} \subset C_{n_i} - p_{a_{n_i}}$, and hence that $C_{n_j} \cdot F_i = 0$ as above, and since F_i is a component of $C_{n_i} - p_{a_{n_i}}$ and C_{n_i} is locally connected, it follows that no point of F_i is a limit point of $\sum_{n=i+1}^{\infty} F_n$; and therefore no point of F_i is a limit point of $F - F_i$, where $F = \sum_1^{\infty} F_n$. But then for each i, F_i is a component of F, which is impossible by virtue of result (4) in § 2, because the diameter of every set F_i is $> \epsilon/2$. Thus the

supposition that $D \neq 0$ for every α leads to a contradiction. Accordingly there exists an α of the first or second class such that $D_\alpha = 0$ and hence such that the set N_α, which is $= N - \sum_{\beta<\alpha} D_\beta$, has no local separating point. Since $N - N_\alpha$ $(= \sum_{\beta<\alpha} D_\beta)$ is countable, therefore N_α is connected.

The set N_α, then, is connected and locally connected, [by (5) in § 2], and has no local separating point. Let A and B be two mutually exclusive open subsets of N_α such that $\bar{A} \cdot \bar{B} = 0$ and hence such that $\rho(A, B) > 0$. Since N_α has no cut point, it follows by Lemma 2 in § 3 that there exist two strongly separated regions R_1 and S_1 in the space N_α each of which joins A and B. Since N_α has no local separating point, S_1 can have no cut point. Thus by Lemma 2 there exist two strongly separated regions R_2 and S_2 in S_1, (S_1 considered as a space), each of which joins A and B, because $A \cdot S_1$ and $B \cdot S_1$ are non-degenerate sets. Likewise S_2 can have no cut point, and hence by the same reasoning it follows that there exist two strongly separated regions R_3 and S_3 in the space S_2 each of which joins A and B, and so on. Continuing this process indefinitely, we obtain an infinite sequence of sets $R_1, R_2, R_3, \cdots$ each of which is a region in the space N_α which joins A and B and such that each pair of these regions are strongly separated. But then for each n, R_n is a component of $\sum_1^\infty R_i$, which is impossible in view of result (4) in § 2, because for each n, R_n joins A and B and hence $\delta(R_n) \geq \rho(A, B)$.

Thus the supposition that our theorem is false leads to a contradiction, and accordingly the theorem is established.

In conclusion it will be noted that our theorem is equally valid for metric continua which are locally compact as for compact metric continua. This is evident at once, because any locally connected and locally compact metric continuum M which is not a rational curve clearly contains compact continuum which is not a rational curve, namely, a closed and compact region in M containing any point p of M in which M is not rational. And if M is hereditarily locally connected, this is impossible by our theorem.

THE JOHNS HOPKINS UNIVERSITY.

GENERALIZATIONS OF BICONNECTED SETS.

By P. M. Swingle.

Introduction.

The problem of the infinite divisibility of space is one which has interested philosophers ever since Zeno stated his paradoxes while geometry was having its origin in ancient Greece. It is a problem which the greatest philosophers of the past centuries gave thought to.

Hume insisted that the mind refused actually to subdivide further after a few divisions.* It seemed to him that the mind was unable to know parts of space, such as present day points, which were obtained by subdivision of a bounded space into more than a finite number of parts. For of such the mind could not have an impression. Kant came a little nearer to present day theory in that he apparently admitted that the mind could continue to subdivide. However in the world of phenomena he did not admit an infinite subdivision as a completed process. But he granted that there might be an unknowable world of noumena in which infinite divisibility existed.†

Thus, in the light of the historical development of mathematics, all kinds of sets, which do not admit of a type of infinite divisibility, are of interest, even though in their definition the type of divisibility objected to is used. Biconnected sets are such sets.‡ For they are connected sets which cannot be subdivided into two distinct § connected subsets. Here infinite divisibility into distinct subsets may exist according to present day mathematics. But infinite divisibility into distinct connected subsets does not exist.

In this paper the various definitions of biconnected sets will be generalized as well as various known theorems concerning them. While these

* David Hume, "A Treatise on Human Nature."

† Immanuel Kant, "Critique of Pure Reason."

‡ For definitions, theorems, and examples of biconnected sets see B. Knaster and C. Kuratowski, "Sur les ensembles connexes," *Fundamenta Mathematicae*, Vol. 2, pp. 206-253. See also J. R. Kline, "A Theorem Concerning Connected Sets," *Fundamenta Mathematicae*, Vol. 3, pp. 238-239. For an interesting example see R. L. Wilder, "A Point Set Which Has no True Quasi-Components and Which Becomes Connected upon the Addition of a Single Point," *Bulletin of the American Mathematical Society*, Vol. 33 (1927), pp. 423-427. For further theorems see also R. L. Wilder, "On the Dispersion Sets of Connected Point Sets," *Fundamenta Mathematicae*, Vol. 4, pp. 214-228. For an unsolved problem see C. Kuratowski, *Fundamenta Mathematicae*, Vol. 3, p. 322 (19).

§ In this paper two sets A and B are said to be *distinct* if $A \times B = 0$.

generalizations will give sets which admit of an infinite divisibility, the main results obtained are for sets which admit only a finite subdivision into distinct connected subsets.

Due to the few known types of biconnected sets and the difficulty of developing the theory of such sets, a number of unsolved problems are stated in this paper in the hope that they will suggest further interesting sets or development of theory.

The results obtained will hold for any space in which the sets exist, unless otherwise stated.

Two Equivalent Generalizations.

Definition. An *n-divisible connected set,* where n is a given cardinal number, is a connected * set which is the sum of n but not of a greater number of distinct connected subsets. Such a set will be said to be *n-divisible.*

Definition. An *n-containing connected set,* where n is a given cardinal number, is a connected set which contains n but not a greater number of distinct connected subsets. Such a set will be said to be *n-containing.*

Examples of *biconnected sets, which are both one-divisible and one-containing,* have been given for the euclidean plane by B. Knaster and C. Kuratowski in their paper *"Sur les ensembles connexes."* † In the example α there given let b be the point $(0,0)$, c the point $(1,0)$, and a the point $(1/2, 1/2)$, where a is the point which totally disconnects the biconnected set. The *notation* (bac), with necessary subscripts, will be used in this paper to denote such a set, wherever a, b, and c may be in the plane. The set (bac) will be understood to contain b, c, and whatever other possible points are desired below.

If $(b_ia_ic_i)$ $(i=1,2)$ are distinct except that $c_1=c_2$, then $(b_1a_1c_1)+(b_2a_2c_2)$ is an example of both a two-divisible and a two-containing connected set. A simple continuous arc would be an example of an n-divisible and n-containing connected set, where n is the power of a countable infinity. Since in a euclidean space the greatest number of points therein contained is the power of the linear continuum, it is seen that such a space does not contain an n-divisible or n-containing set, where n is greater than the power of the linear continuum. But if n is the power of the linear continuum, a euclidean

* A set M will be said to be *connected* if it contains at least two points and for every two distinct non-vacuous subsets of M, whose sum is M, at least one of these contains a limit point of the other. By this definition a point will not be considered connected.

† *Loc. cit.*

space of dimensions greater than one is itself both an n-divisible and an n-containing connected set.*

LEMMA 1. *If the connected set M contains n distinct connected subsets, (N), where n is a given cardinal number, then, for any positive integer w not greater than n, M is the sum of w distinct connected subsets, (C), such that each set of (C) contains at least one set of (N).*

Let $N_1, N_2, \cdots, N_w$ be w of the sets of (N). Let C_1 be the maximal connected subset of $M-(N_2+N_3+\cdots+N_w)=Z$ which contains N_1. Let $M-C_1=M_1+M_2+\cdots+M_k$ separate.† It is necessary that k be less than w. For if not there exists an M_i, M_h say, which does not contain a point of $M-Z$, and so, as M_h+C_1 is connected, C_1 is not a maximal connected subset of Z. It is necessary then that each of the sets M_i $(i=1,2,\cdots,k)$, where k has its maximum value, be connected and contain a point of an N_g $(g=2,\cdots,w)$. Let C_1 and each M_i, which contains one and only one N_g, be each a set of (C). The remaining sets M_i can now be treated as M was above. Thus the sets of (C) are obtained.

The truth of the following corollary is now evident, giving a type of an "any to finite" property.

COROLLARY 1. *If the connected set M contains infinitely many distinct connected subsets, then M is the sum of w distinct connected subsets, where w is any positive integer.*

That lemma 1 does not hold if both n and w are the power of a countable infinity is seen from the following example. Let $(b_ia_ic_i)$ $(i=1,2,\cdots)$ be a countable infinity of biconnected sets, which are distinct except that $(b_ia_ic_i)\times(b_{i-1}a_{i-1}c_{i-1})=c_i=b_{i-1}$ for every i; let these biconnected sets have the further property that they have a simple continuous arc t as sequential limiting set, which has nothing common with any $(b_ia_ic_i)$. Let q be any point of t. The set $(b_1a_1c_1)+(b_2a_2c_2)+\cdots+q$ will be called *a set* (bq) in this paper. The set (bq) is an n-containing connected set, where n is a countable infinity, which contains the n distinct biconnected sets $(b_ia_ic_i)-b_i$. However (bq) is not n-divisible as there do not exist n distinct connected subsets of M, one of which contains q, of which M is the sum. This set is an example of a set defined as follows.

* For a space containing more elements than the power of the linear continuum see F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig 1914, p. 68. For other spaces see also pp. 284-290.

† By the notation $M_1+M_2+\cdots+M_k$ *separate* is meant that the sets M_i $(i=1,2,\ldots,k)$ are distinct, non-vacuous, sets, no one of which contains a limit point of the sum of the remaining ones.

Definition. A connected set M is said to be *finitely-divisible* if it is neither the sum of a maximum finite number nor of an infinite number of distinct connected subsets.

Problem 1. It is not determined in this paper if the first part of lemma 1 holds if n is the power of the linear continuum and if w is either the power of a countable infinity or $w = n$.

THEOREM 1. *In order that the connected set M be n-divisible, where n is a positive integer, it is necessary and sufficient that M be n-containing.**

The condition is necessary. For as M is n-divisible it contains n distinct connected subsets. Hence, if it is not n-containing, it must contain more than n, and so $n + 1$, such subsets. Thus by lemma 1 it cannot be n-divisible.

The condition is sufficient. For, if M is n-containing, it is by lemma 1 the sum of n distinct connected subsets. And, if it is the sum of more than n such subsets, it is not n-containing.

COROLLARY 2. *If M is an n-divisible connected set lying in a locally compact metric space, where n is any positive integer, then M is punctiform.*†

For if M is not punctiform it contains a subcontinuum W. And in a locally compact metric space any vicinity which contains a point of W contains a subcontinuum of W. Therefore W must contain $n + 1$ distinct connected subsets and so, by theorem 1, M is not n-divisible.

THEOREM 2. *If M is an n-divisible connected set, where n is any positive integer, then M is the sum of n distinct biconnected subsets.*

As M is n-divisible it is the sum of n distinct connected subsets, no one of which can be the sum of two distinct connected subsets, as M cannot be the sum of $n + 1$ distinct connected subsets. Hence each of the distinct connected subsets, of which M is the sum, must be biconnected.

It is of interest to note that an n-divisible connected set may be the sum of less than n distinct biconnected subsets. For consider the three-divisible connected set $(b_1a_1c_1) + (b_2a_2c_2) + (b_3a_3c_3)$, where $b_1c_1 = b_2c_2$, $(b_2a_2c_2)$ is obtained by rotating $(b_1a_1c_1)$ $180°$ about b_1c_1, $a_3c_3 = b_2a_2$, and otherwise the three biconnected sets are distinct. Then this three-divisible connected set is the sum of the two distinct biconnected subsets $(b_3a_3c_3)$ and $[(b_1a_1c_1) + (b_2a_2c_2)] - b_2a_2$.

* For a proof of this theorem for $n = 1$ see B. Knaster and C. Kuratowski, *loc. cit.*, p. 215, theorem 11.

† For a proof of this corollary for $n = 1$ see B. Knaster and C. Kuratowski, *loc. cit.*, p. 216, theorem 14.

THEOREM 3. *If M is an n-divisible connected set, where n is any positive integer, then M contains at most $2n-1$ points each of which disconnects M.**

As M is n-divisible it is the sum of n distinct biconnected subsets by theorem 2. Let C_i $(i=1,2,\cdots,n)$ be these n sets. Hence each point q that disconnects M is contained in one and only one set C_i. And either q disconnects C_i or it does not. Of those points which disconnect both M and a C_i there can be at most n, since there exists at most one point which disconnects a biconnected set.

Consider now the set K, each point of which disconnects M but does not disconnect the C_i which contains it. Consider all points of K that are limit points of C_1. They are of three classes, which we now proceed to consider.

Let p_1 be a point of K in C_1. Then $M-p_1=M_1+M_2$ separate, where M_1 contains the connected set C_1-p_1 say. Now any set C_i which has a point in M_2 lies wholly in M_2. Let the class of those sets C_i that lie in M_2 and have p_1 as a limit point be denoted by G_1. No set C_i in G_1 has any other limit point then p_1 in C_1, nor does it contain a limit point of C_1, as M_1+M_2 separate. Let p_2 be another point of K in C_1. Then $M-p_2=H_1+H_2$ separate, where C_1-p_2 lies in H_1 say. As before denote by G_2 the class of those sets C_i in H_2 that have p_2 as a limit point. Since the observation made above concerning the sets of G_1 apply also to the sets of G_2, no set of G_1 lies also in G_2. As the number of sets C_i is finite, we proceed, in this manner, to a finite set of points $p_1, p_2, \cdots, p_{k_1}$ with which are associated, respectively, classes $G_1, G_2, \cdots G_{k_1}$ of sets C_i.

Let x_1 be a point of K in some C_h that is a limit point of C_1 and separates C_1 from C_h in M. Obviously C_h is not in any G_i, and C_h has no other limit point of C_1, nor does it have a limit point in C_1. Let F_1 be the set of all those sets C_i that have x_1 as a limit point, and are separated from C_1 by it. Proceed in this manner to get sets $F_2, F_3, \cdots, F_{k_2}$ corresponding to points $x_2, x_3, \cdots, x_{k_2}$, respectively. As before we see that the classes $F_1, F_2, \cdots, F_{k_2}$ are distinct, and are distinct from the G's. As every point of K that lies in C_1 is a point of the set $p_1, \cdots, p_{k_1}$, and every point that is a point of K and a limit point of C_1 and separates C_1 from the C_i containing it is a point of the set $x_1, \cdots, x_{k_2}$, any other point of K must be a point having neither of these properties. There may still be for instance a point that is a limit point of C_1 and in a C_j, but does not separate C_1 and C_j. Such points we consider next.

* For the case where $n=1$ see J. R. Kline, *loc. cit.*; also see C. Kuratowski, *loc. cit.*

Let y_1 be such a point. It separates C_j from some C_k. Then with y_1 associate all sets C_i such as C_k; i. e., sets which have y_1 as a limit point and are separated from C_j by y_1; denote the class of such sets by T_1. No set of T_1 is in an F or a G. That it is not in a G is easy to see. If it were in an F, it would be separated from C_1 by a point x_m, which lies in a $C_n \neq C_j$ and hence is distinct from y_1. This is impossible.

Thus with every point of K that is a limit point of C_1 we associate at least one set C_i and in such a way that there is no overlapping between the sets associated with different points. Let Z_1 denote the point set consisting of all points in C_1, and all points in G's, F's, and T's, together with the C's which have limit points in these sets which are not contained in K.

To complete the proof we note that no point of $Z_1 + C_1$ that is in K separates $Z_1 + C_1$, and we can proceed from this set as we proceeded from C_1. In this manner we associate with each point of K at least one set $C_i \neq C_1$ and in such a way that overlapping is avoided. Consequently the number of points in K is less than or equal to $n - 1$.

Hence there exist at most $n + (n - 1) = 2n - 1$ points each of which disconnects M.

THEOREM 4. *If M is an n-divisible connected set, where n is a positive integer, then no connected subset C of M is irreducible connected about any subset N of C such that $C - N$ contains more than $2n - 1$ points.**

By theorem 1 there exists an integer q, not greater than n, such that C is q-containing. Let N be any set about which C is irreducibly connected. Then any point p of $C - N$ disconnects C otherwise $C - p$ is a proper connected subset of C containing N. Hence $C - N$ can contain at most $2n - 1$ points by theorem 3.

ANOTHER GENERALIZATION.

Definition. Let $1, 2, \cdots; \omega, \omega + 1, \omega + 2, \cdots; 2\omega, 2\omega + 1, \cdots; \cdots; \cdots, k$ be the set of the first k ordinal numbers and let $W_1, W_2 \cdots W_k$ be a set of k connected subsets of the connected set W. Then $W_1, W_2, \cdots, W_k$ is called a *k-convergable sequence of W* when $W_1 = W$, W_g contains W_i, where i runs over all ordinal numbers not greater than k and g over all less than i, and $W_g - W_{g+1}$ is a connected subset of W, but for every connected subset C of W_k, $W_k - C$ is either a point, a vacuous, or a totally disconnected point set.

---

* For the case where $n = 1$ see B. Knaster and C. Kuratowski, *loc. cit.*, p. 225. theorem 29.

Definition. A connected set W is an *n-convergable connected set,* where n is any cardinal number, if there exists an ordinal number k, the cardinal number of which is n, such that W contains a k-convergable sequence, but there does not exist an ordinal number q, the cardinal number of which is greater than n, such that W contains a q-convergable sequence. The set W will be called also *n-convergable.*

A biconnected set is an example of a one-convergable set.*

Definition. A connected set W is a *finitely-convergable connected set* or is *finitely-convergable* if it contains, for every positive integer N, an n-convergable sequence such that n is an integer $\geqq N$, but it does not contain a q-convergable sequence, where q is an infinite cardinal number.

The example (bq), given above, of a finitely-divisible connected set is also an example of a finitely-convergable set. However $(bq) - q$ is an example of a finitely-convergable but not of a finitely-divisible connected set.

The example $(bac) + bd$, where bd is an arc distinct from (bac) except for the point b, is an example of an n-convergable set, where n is the power of a countable infinity.

In the euclidean plane S let (t) be the set of line segments from $(i, 0)$ to $(i, 1)$, where i takes on all real values from zero to one. Let the sets of (t) be well ordered by Zermelo's postulate, obtaining the set (t_i). And let (bac) be a biconnected set which has no point common with any of the sets of (t). Let T be the set of points of S contained in neither (bac) nor in a set of (t) and let $(t)_j$ be the points of S contained in the first j sets of (t_i). Then there exists an ordinal number $k + 1$, whose cardinal number n is the power of the linear continuum, such that $W_1 = S$, $W_2 = S - (t)_1$, $W_3 = S - (t)_2, \cdots;$ $\cdots; \cdots, W_{j+1} = S - (t)_j \cdots; \cdots, W_k = T + (bac)$, and $W_{k+1} = (bac)$. As all the W_i's except W_{k+1} contain T it is seen that they are connected. And $W_i - W_{i+1} = t_i$ is connected as is also $W_k - W_{k+1} = T$. Hence $W_1, W_2, \cdots$, W_k, W_{k+1} is a $(k + 1)$-convergable sequence of S. And as the power of the set of distinct points of S is the power c of the linear continuum, there does not exist an ordinal number q, whose power is greater than c, such that S contains a q-convergable sequence. Hence S, and similarly any euclidean space of dimension > 1, is a c-convergable connected set, where c is the power of the linear continum.

THEOREM 5. *If W is an n-convergable connected set, where n is a positive integer, then W is the sum of n distinct biconnected subsets.*

As W is n-convergable it contains an n-convergable sequence $W_1 = W$,

* B. Knaster and C. Kuratowski, *loc. cit.*, p. 215, theorem 11.

$W_2, \cdots, W_n$, where it is known that W_n is biconnected. Let $W_i - W_{i+1} = V_{n-i+1}$ $(i = 1, 2, \cdots, n-1)$ and let $W_n = V_1$. Let j be such that $W_{n-j+1} = V_1 + V_2 + \cdots + V_j$, where each of the sets $V_1, V_2, \cdots, V_j$ is a biconnected set. Assume that V_{j+1} is not biconnected and so is the sum of the two distinct connected subsets X and Y. Then as $W_{n-j} = W_{n-j+1} + V_{j+1} = W_{n-j+1} + (X + Y)$, either $X + W_{n-j+1}$ or $Y + W_{n-j+1}$ is connected. Take for example the case where $Z_2 = Y + W_{n-j+1}$ is connected. Let $Z_1 = W_{n-j}$ and $Z_g = W_{n-j+g-2}$ $(g = 3, 4, \cdots, j+2)$. Hence $Z_1, Z_2, \cdots, Z_{j+2}$ is a $(j+2)$-convergable sequence of the $(j+1)$-convergable connected set W_{n-j} which is a contradiction. It is then necessary that W be the sum of the n distinct biconnected sets $V_1, V_2, \cdots, V_n$.

The example given of an n-convergable set, where n is the power of a countable infinity shows that theorem 5 is untrue for such a set. And the following example shows that it is untrue for the case where n is the power of the linear continuum.

Example A. In the euclidean plane consider the straight line interval g from $(0, 0)$ to $(1, 0)$ and the set of line segments (t) from $(i, 0)$ to $(i, 1)$ where i takes on the irrational values from zero to one. Let $T = g + (t)$ and let (bac) have but the point b common with T where g contains b. It is seen that $T + (bac)$ is a c-convergable connected set, where c is the power of the linear continuum. It is seen further that T does not contain a biconnected set V, for V would have to contain more than one point of g and so there would exist more than one point which disconnects the biconnected set V, which is impossible. Hence it follows that $T + (bac)$ cannot be the sum of c distinct biconnected subsets.

Problem 2. Does there exist a cardinal number q such that a euclidean space is the sum of q distinct biconnected subsets? Such a space is the sum of c distinct indivisible subsets, that is points, where c is the power of the linear continuum.

LEMMA 2. *If the connected set W is the sum of the k, k a positive integer, distinct connected subsets $C_1, C_2, \cdots, C_k$, then there exist k connected subsets $W_1 = W, W_2, \cdots, W_k = C_k$, where W_i contains W_{i+1} $(i = 1, 2, \cdots, k-1)$ and $W_i - W_{i+1}$ is connected.*

As W is connected, either $W_k = C_k$ has a limit point in one of the sets $C_1, C_2, \cdots, C_{k-1}$ or one of these sets has a limit point in W_k. Hence let one of these sets, which has either of these properties, together with the set W_k form the set W_{k-1}. And now either W_{k-1} has a limit point in one of the

remaining sets of $C_1, C_2, \cdots, C_{k-1}$ or one of these sets has a limit point in W_{k-1}. Let one of these sets, having either property, together with W_{k-1} form the set W_{k-2}. Proceeding in this manner the truth of the lemma is seen.

THEOREM 6. *If W is an n-divisible connected set, where n is a positive integer, then W is an n-convergable connected set.*

As W is n-divisible, by theorem 2 it is the sum of n distinct biconnected subsets $C_1, C_2, \cdots, C_n$. Hence by lemma 2 it is seen that W contains an n-convergable sequence $W_1 = W$, $W_2, \cdots, W_n = C_n$, as C_n is biconnected. And as an n-divisible connected set contains at most n distinct connected sets, W does not contain a q-convergable sequence, where q is greater than n. Thus W is n-convergable.

As a simple continuous arc is n-divisible, where n is the power of a countable infinity, but is not n-convergable, it is seen that theorem 6 does not hold for such an n. And in Example A the set T is c-divisible but not c-convergable, where c is the power of the linear continuum.

Problem 3. Does there exist a finitely-divisible connected set which is not finitely-convergable? An example has been given above of a finitely-convergable set which is not finitely-divisible.

A theorem will be proven now for finitely-convergable connected sets which corresponds to theorem 5 for n-convergable sets, n finite.

THEOREM 7. *Let W be a finitely-convergable set. Then either W contains an infinite sequence of distinct biconnected sets $\cdots, M_3, M_2, M_1$ or for every integer k there exists an integer q greater than or equal to k such that W is the sum of the q distinct biconnected sets $M_q, M_{q-1}, \cdots, M_2, M_1$. The set $M_1 + M_2 + \cdots + M_t = H_t$ $(t = 1, 2, \cdots)$ may be taken connected and $W - H_t$ the sum of a finite number of maximal connected subsets such that, if C is one of these, then $H_t + C$ also is either the sum of a finite number or contains an infinite number of biconnected sets. And the set $W - H_\infty$ does not contain a biconnected set which for every finite t, is contained in the maximal subset of $W - H_t$ which contains M_{t+1}.*

For any integer k there exists an integer n, greater than k, such that W contains an n-convergable sequence $W_1, W_2, \cdots, W_n$. Then, as in the proof of theorem 5, we obtain the sets $V_1, V_2, \cdots, V_n$, of which the first is known to be biconnected, and if it is assumed that the first j of these sets is biconnected but that V_{j+1} is not, one obtains the $(j+2)$-convergable sequence of W_{n-j} obtained in theorem 5. Let this sequence be $Z_1 = W_{n-j}$, $G_1 = W_{n-j+1} + Y$, $Z_2 = W_{n-j+1}$, $Z_3 = W_{n-j+2}, \cdots, Z_{j+1} = W_n$. If now G_1

$-Z_2$ is not a biconnected set, it can be treated as V_{j+1} was above, obtaining a $(j+3)$-convergable sequence $Z_1, G_1, G_2, Z_2, Z_3, \cdots, Z_{j+1}$. Proceeding in this manner it is seen that either there exists an integer h such that there exists a $(j+h+1)$-convergable sequence $Z_1, G_1, G_2, \cdots, G_h, Z_2, Z_3, \cdots, Z_{j+1}$, such that $G_h - Z_2$ is biconnected, where h may equal one, or one obtains the m-convergable sequence, where m is the power of a countable infinity, $W_1, W_2, \cdots, W_{n-j} = Z_1, G_1, G_2, \cdots; V_j$. As the latter is contrary to the fact that W is finitely-convergable, there exists the $(j+h+1)$-convergable sequence of which $G_h - Z_2 = M_{j+1}$ is biconnected. Hence by mathematical induction it follows that either W must be the sum of q distinct biconnected subsets, where q is greater than or equal to k, or W contains the infinite sequence of distinct biconnected subsets $\cdots, M_3 = V_3, M_2 = V_2, M_1 = V_1$. From the derivation it is seen that H_t, t finite, is connected and $W - H_t$ is always the sum of a finite number of maximal connected subsets, of which, if C is one, $H_t + C$ is either the sum of a finite number or contains an infinite number of distinct biconnected subsets. And $W - H_\infty$ does not contain a biconnected subset B which for every finite t, is contained in the maximal connected subset of $W - H_t$ which contains M_{t+1}, for $W_1, W_2 = W - H_1, \cdots; B$ or a similar sequence, is an m-convergable sequence, where m is the power of a countable infinity, which is impossible.

This theorem suggests the existence of the following interesting and more complicated example of a finitely-convergable set. Let (bq) be a set similar to the example of a finitely-convergable set given above. Let (b_iq_i) $(i=1, 2, \cdots, 6)$ be six such sets distinct except that $(b_1q_1)\times(b_2q_2)$, $(b_2q_2)\times(b_3q_3)$, $(b_2q_2)\times(b_4q_4)$, $(b_4q_4)\times(b_5q_5)$, and $(b_1q_1)\times(b_6q_6)$ say each contain one and only one point, which is neither a q_i nor is it a point which totally disconnects one of the biconnected subsets. Then $(b_1q_1)+(b_2q_2)+\cdots+(b_6q_6)$ is a finitely-convergable set. The theorem further suggests the following problem.

Problem 4. Is a finitely-convergable connected set ever the sum of a finite number of distinct biconnected subsets?

The following example of an n-convergable set, where n is the power of a countable infinity, is of interest in this connection.

Example B. Let $(baca')$ be the biconnected set (bac) together with the biconnected set $(ba'c)$ obtained by rotating (bac) 180° about bc. Let $(b_ia_ic_ia_i')$ $(i; 1, 2, \cdots)$ be an infinite number of such sets, whose sum is bounded, which are distinct except that $a_1 = a_2 = a_3 = \cdots$, and let (bac) be a biconnected set such that bc of (bac) contains $a_1' + a_2' + \cdots$, but (bac) contains nothing else common with the sets $(b_ia_ic_ia_i')$. Then $W = (b_1a_1c_1a_1') + (b_2a_2c_2a_2') + \cdots + (bac)$ is both an n-convergable and an

n-divisible connected set, where n is the power of a countable infinity. It has the interesting property that it is the sum of q distinct biconnected subsets, where q is either any integer $\geqq 2$ or is a countable infinity.

Example C. In Example B let (bac) have the further property that $a = a_1$. The resulting set W is still both n-divisible and n-convergable and is the sum of a countable infinity of distinct biconnected subsets. However it is no longer the sum of a finite number of distinct biconnected subsets.

These examples suggest the following problems.

Problem 5. If for every integer n, greater than one, the connected set M is the sum of n distinct biconnected subsets, is M then the sum of infinitely many such distinct subsets? In previous theorems conditions have been given which cause a connected set to be the sum of a finite number of distinct biconnected subsets. Here it is asked what conditions cause a connected set to be the sum of infinitely many such subsets.

Problem 6. If the connected set M is the sum of a finite number of distinct biconnected subsets but is not the sum of an infinite number of distinct connected subsets, does there exist a finite n such that M is n-divisible? This problem is of interest in connection with theorem 2.

Problem 7. If the connected set M does not contain a maximum finite number of distinct connected subsets, must it contain an infinite number of such subsets, i. e., does there exist a *finitely-containing connected set*?

n-Convergable Sets.

A number of other theorems will now be proven concerning n-convergable sets, where n is a positive integer. In theorem 6 it was shown that if a connected set is n-divisible it is n-convergable. Thus we have the following problem.

Problem 8. Does there exist an n-convergable connected set, where n is a positive integer, which is not n-divisible?

The following theorems will be of interest in connection with this and other problems.

Theorem A. *Let W be an n-convergable connected set, where n is a positive integer, which is not also an n-divisible set, if such a set exists. Let C_i $(i = 1, 2, \cdots, n)$ be the n distinct biconnected subsets of which W is the sum and let N_j $(j = 1, 2, \cdots, n + k)$ be a finite number, greater than n, of distinct connected subsets of which W is the sum. Let C be any C_i and N be any N_j. Then* (1) *N is not biconnected;* (2) *if M is any connected*

subset of W which is the sum of more than n distinct connected subsets, as N is, then $C \times M \neq 0$; (3) *$C - N \times C$ is totally disconnected; and* (4) *$N - C \times N$ is totally disconnected if it contains more than one point.*

It follows at once that (1) is true, since by lemma 2 W would contain an $(n+k)$-convergable sequence otherwise.

Assume that $C \times M$, in (2), is vacuous. As $W - C$ is the sum of a finite number of maximal connected subsets, one of these, Z say, contains M. Hence by lemma 1 Z is the sum of more than n distinct connected subsets. Thus it follows that W is also, and one of these distinct connected subsets is the biconnected set C, which is a contradiction according to (1). It also follows from (1) that N is a set such as M.

Assume that $C - N \times C = (N + C) - N$ contains the maximal connected subset K, which must be biconnected as C contains it. Then $(N + C) - K$ is connected and contains N, and, as by (2) N contains a point of every C_i, $W - K$ must be connected. Since $W - K$ contains N, by lemma 1 it is the sum of more than n distinct connected subsets, and so a contradiction with (1) is obtained as K is biconnected.

Assume that $N - C \times N = (N + C) - C$ contains the maximal connected subset X. Then by (1) X cannot be the sum of a finite number of distinct connected subsets, one of which is biconnected; for $W - X$, which contains C and the connected subset $(N + C) - X$, is connected, since by (2) C contains a point of every N_i; and $W - X$ is the sum of more than n distinct connected subsets by lemma 1, since it contains an N_i. Hence X must be the sum of more than n distinct connected subsets. But this is impossible by (2) as $X \times C = 0$.

LEMMA A. *If W is an n-convergable connected set, where n is a positive integer, but W is not n-divisible, then there does not exist a finite subset which disconnects W.*

Assume that the set Q, which contains q points, disconnects W. Let g be an integer greater than both n and q. Hence as W is not n-divisible it is necessary by theorem A (1) that W be the sum of g distinct connected subsets N_i $(i = 1, 2, \cdots, g)$. Thus one of the N_i's, N say, does not contain a point of Q. Let $W - Q = M_1 + M_2$ separate, where M_1 contains N. As by theorem A, N contains a point of each biconnected set C_j $(j = 1, 2, \cdots, n)$, of which W is the sum, M_1 does also. Let C be a C_j which contains also points of M_2. Hence $C - C \times Q = K_1 + K_2$ separate, where M_1 contains K_1 and M_2 contains K_2. As $C - N \times C$, which is totally disconnected by theorem A, contains $K_2 + Q \times C$, $K_2 + Q \times C = Z_1 + Z_2 + \cdots + Z_g$ sepa-

rate, where $Z_g \times Q = 0$. Hence $C = (K_1 + Z_1 + Z_2 + \cdots + Z_{g-1}) + Z_g$ separate, which is a contradiction. Thus no finite subset disconnects W.

THEOREM 8. *If W is an n-convergable connected set, where n is a positive integer, then W contains at most $2n - 1$ points each of which disconnect W.*

If W is not n-divisible the truth of the theorem follows from lemma A. And if it is n-divisible it follows from theorem 3.

The truth of the following corollary is now evident.

COROLLARY 3. *If W is an n-convergable connected set, where n is a positive integer, then W is not irreducibly connected about any set N such that $W - N$ contains more than $2n - 1$ points.*

In the previous theorems on n-convergable connected sets, where n is a positive integer, the full power of the definition of such sets has not been used. Only properties have been made use of which are given by the following definition.

Definition. A connected set W will be said to be *n-convergable on a finite range,* where n is a positive integer, if W contains an n-convergable sequence but does not contain an $(n+1)$-convergable sequence.

The theorems proved this far for n-convergable connected sets hold for all sets n-convergable on a finite range. It is evident that an n-convergable set is n-convergable on a finite range. But we have the following problem, since a set n-convergable on a finite range might contain a q-convergable sequence, where the power of q is a transfinite cardinal number.

Problem 9. Does there exist a set n-convergable on a finite range which is not n-convergable?

Two theorems will now be proven which use in their proof more than is given apparently in the definition of sets n-convergable on a finite range.

THEOREM 9. *If B is a biconnected subset of an n-convergable connected set W, where n is a positive integer, then W is the sum of a finite number, less than or equal to n, of distinct biconnected subsets, a number of which form a connected set Z containing B, such that $Z - B$ is either vacuous, a point, or a totally disconnected set.*

By theorem 5 it is seen that the theorem is true unless $W - B$ contains a maximal connected subset.

Assume then that $W - B$ contains the maximal connected subset T. Then W is the sum of the two distinct connected subsets T and $W - T$, the latter of which contains B. Consider for example the case where T is not the sum of a finite number of distinct connected subsets one of which is bi-

connected. Then T is the sum of two distinct connected subsets U_1 and V_1. Consider for example the case where $U_1 + (W - T)$ is connected, since if it is not, $V_1 + (W - T)$ is. Also U_1 is the sum of two distinct connected subsets U_2 and V_2, where U_2 is such say that $U_2 + (W - T)$ is connected. Proceeding in this manner one obtains the k-convergable sequence, where k is a countable infinity, $W_1 = W$, $W_2 = U_1 + (W - T)$, $W_3 = U_2 + (W - T)$, $\cdots$; B. As this is a contradiction it follows that T, and likewise the set composed of all maximal connected subsets of $W - B$, must be the sum of a finite number, less than n, of distinct biconnected subsets. Let (B) represent the sum of these distinct biconnected subsets of $W - B$. Hence $W - (B) = Z$ is connected and so is composed of a finite number of distinct biconnected subsets. And it is evident now that $Z - B$ is either vacuous, a point, or a totally disconnected point set.

THEOREM 10. *Let W be an n-convergable connected set, where n is a positive integer. Let B be a biconnected and C a connected subset of W. Then* (1) *no connected subset of W contains more than n distinct connected subsets where one of them contains a biconnected subset;* (2) *either C is the sum of not more than $n - 1$ distinct connected subsets or $C \times B \neq 0$;* (3) *either C and $C + B$ are the sum of not more than n distinct biconnected subsets or $B - B \times C$ is vacuous or totally disconnected;* (4) *either $(C + B)$ is the sum of not more than n distinct connected subsets or $C - B \times C$ is totally disconnected; and* (5) *if C is not the sum of distinct biconnected subsets then $C - C \times B$ contains at most one biconnected subset F and $C - F$ is totally disconnected.*

Assume that W contains the connected subset K which is the sum of more than n distinct connected subsets one of which contains the biconnected subset E. Then by lemma 1, W is the sum of more than n distinct connected subsets, $U_1, U_2, \cdots, U_g$, where U_g contains E. Hence there exists, by lemma 2, the sequence of connected sets $W_1 = W$, W_2, $\cdots$, $W_g = U_g$, where $W_i - W_i + 1$ $(i = 1, 2, \cdots, g - 1)$ is connected, but by (1) of theorem A it is not biconnected. Thus the k-convergable sequence $W_1, W_2, \cdots, W_{g-1}$, $Z_1, Z_2, \cdots$; E is obtained, where k is a countable infinity, $W_{g-1} - W_g$ contains $Z_j - Z_{j+1}$ $(j = 1, 2, \cdots)$ and Z_j contains W_g. As this is a contradiction, (1) is true.

Assume that C is the sum of at least n distinct connected subsets and that $C \times B = 0$. Then by lemma 1, W is itself the sum of more than n distinct connected subsets, one of which contains the biconnected set B. As this is contrary to (1) it is seen that (2) is true. Also it is seen that if $C \times B = 0$, $C + B$ is the sum of not more than n distinct connected subsets.

Assume that $B - B \times C = (B + C) - C$ contains the maximal connected, and so biconnected, subset E. Then, as $E \times C = 0$, by (2) C is the sum of not more than $n - 1$ distinct connected subsets. And as $(B + C) - E$ is connected, and has nothing common with the biconnected set E, by (2) it must be the sum of not more than $n - 1$ distinct connected subsets. Thus by (1) $B + C$ is the sum of not more than n distinct biconnected subsets. Hence (3) is true.

Assume that $C - B \times C = (B + C) - B$ contains the maximal connected subset F. Then, as $F \times B = 0$, by (2) F is the sum of not more than $n - 1$ distinct connected subsets and so of less than n distinct biconnected subsets. Similarly $(B + C) - F$ is the sum of distinct biconnected subsets and so $B + C$ is the sum of not more than n distinct biconnected subsets by (1). Hence (4) is true.

Assume that C is not the sum of distinct biconnected subsets and that $C - B \times C$ contains the biconnected subset F. Then by (4), as $C + F = C$, $C - F$ is totally disconnected. Hence (5) is true.

Problem 10. Does an n-convergable connected set, where n is a positive integer, contain a connected subset which contains no biconnected subset?

Other Theorems.

Theorem 11. *Any connected set M, in a locally compact metric space, which contains both a subcontinuum C and a biconnected subset B which is locally connected at a point q, contains also a k-convergable sequence, where k is a countable infinity.*

There exists a region R_1 containing q, and so containing a biconnected subset E, and a region R_2 containing a point of C such that $R_1' \times R_2' = 0$ and $R_2' \times C$ contains a subcontinuum K. There exists a proper subcontinuum K_1 of K such that $K - K_1$ contains a subcontinuum. Let T_1 be a maximal connected subset of $M - K_1$ which contains a subcontinuum of $K - K_1$. Then $M - T_1$ is connected and contains K_1. Hence M is the sum of the two distinct connected subsets T_1 and $M - T_1$ one of which, T_1 say, contains E and a subcontinuum C_1. Proceeding as above it can be shown that T_1 is the sum of two distinct connected subsets T_2 and $T_1 - T_2$, one of which, T_2 say, contains E and a subcontinuum C_2. Thus it can be shown that M contains the k-convergable sequence $W_1 = M$, $W_2 = T_1$, $W_3 = T_2, \cdots; E$.

Corollary 4. *If a finitely-convergable set M, in a locally compact metric space, contains a biconnected subset which is locally connected at a point, then M is punctiform.*

Problem 11. Is a finitely-convergable or finitely-divisible set, in a locally compact metric space, always punctiform?

LEMMA 3. *Any connected set which is the sum of a finite number of distinct connected subsets, one of which is biconnected, is not irreducibly connected about a finite point set.*

Assume that the connected set M is irreducibly connected about the finite subset H. But as a biconnected set is disconnected by at most one point there exists in the biconnected subset of M a point q, not contained in H, such that $M - q$ is connected. Hence the lemma is true.

COROLLARY 5. *A finitely-convergable set M is not irreducibly connected about a finite subset.*

This follows at once from lemma 3.

LEMMA 4. *Any connected set M is either the sum of a finite number of distinct connected subsets, one of which is biconnected, or it contains a connected subset which is the sum of a countable infinity of distinct connected subsets.*

If M is not the sum of a finite number of distinct connected subsets, one of which is biconnected, it is the sum of two distinct connected subsets U_1 and V_1. And U_1 is the sum of two distinct connected subsets U_2 and V_2, where V_2 is such say that $V_1 + V_2$ is connected since if it is not $V_1 + U_2$ must be. Proceeding in this manner it is seen that the theorem is true.

THEOREM 12. *A finitely-divisible connected set M is not irreducibly connected about a finite subset.*

Assume that M is irreducibly connected about the finite subset Q, which contains q points. Then by lemmas 3 and 4 it is seen that M contains a connected subset H which is the sum of a countable infinity of distinct connected subsets. Then $M - H = M_1 + \cdots + M_{q+1}$ separate, as M is not the sum of a countable infinity of distinct connected subsets. Say for example that $M_1 + M_2 + \cdots + M_q$ contains $Q - H \times Q$. Then $H + M_1 + M_2 + \cdots + M_q$ is a proper connected subset of M containing Q. As this is impossible under our assumption, it is seen that the theorem must be true.

OHIO STATE UNIVERSITY.

SOME METRIC PROPERTIES OF DESCRIPTIVE PLANES.

By J. L. Dorroh.

Introduction.

This paper is concerned with the study of certain properties of planes in which Axioms I-VIII of Veblen's *System of Axioms for Geometry* * are satisfied. Such a plane will be called a *descriptive plane.* A point P of a descriptive plane S will be said to be a limit point of a subset L of S if and only if every triangle of S which incloses † P also incloses a point of L distinct from P.

From the results obtained in Part I, it may be concluded that *a necessary and sufficient condition that a descriptive plane S be metric is that S contain a countable set of distinct points which has a limit point.* In Part IV, it is shown that *not every descriptive plane is metric.*

A conclusion which may be drawn from Part II is that *a necessary and sufficient condition that a descriptive plane S be in one-to-one continuous correspondence with an everywhere dense subset of the number-plane is that S contain a separable segment.* In Part III, it is shown that *not every metric descriptive plane is separable.*

Part I.

Let S denote a descriptive plane which contains a countable set of distinct points which has a limit point. The first of the following theorems can be established by means of projections.

* O. Veblen, "A System of Axioms for Geometry," *Transactions of the American Mathematical Society*, Vol. 5 (1904), pp. 343-384.

† A polygon will be said to inclose a point P if its interior contains P. The term "polygon" is used, in this paper, in the sense of "simple polygon" as defined by Veblen, *loc. cit.*, p. 363, Definition 9. For a definition of the interior of a triangle, see Veblen's Definition 5, *ibid.*, p. 345. A simple polygon separates the plane into just two domains (Veblen's Theorem 28, *loc. cit.*); if K is a polygon which separates the plane into the domains D_1 and D_2, just one of these domains—say D_1—is a subset of the sum of the interiors of a finite number of triangles; D_1 will be called the interior of K, and D_2 will be called the exterior of K.

THEOREM 1. *If A and B are two distinct points of S, there exists a sequence of points $P_1, P_2, P_3, \cdots$, such that:*

(1) AP_iB* $(i = 1, 2, 3, \cdots)$.

(2) $AP_{i+1}P_i$.

(3) A is the sequential limit point of the P_i.

THEOREM 2. *S is metric.*

In order to establish Theorem 2 a method for defining distance in S will be indicated.

Let A, B, C, D, Q denote distinct points of S such that A, C, D are non-collinear, ABC and DCQ. Let $c_1, c_2, c_3, \cdots$ denote the positive rational fractions less than or equal to $\frac{1}{2}$ ($c_1 = \frac{1}{2}$, and $c_i \neq c_j$ if $i \neq j$) which in their lowest terms have for their denominators integral powers of 2. For each positive integer i a point p_i will be selected to correspond to c_i, the p_i being selected so that:

(1) Ap_iQ.

(2) *If* $c_i < c_j$, then Ap_ip_j.

(3) If c_k is the lower limit of a subset $[c]_k$ of the c_i, then p_k is a limit point of the subset of the p_i which correspond to the fractions in $[c]_k$.

(4) A is a limit point of the p_i.

A method for so selecting the p_i will now be described. Let p_1 denote a point in the order Ap_1Q. Let $g_1, g_2, g_3, \cdots$ denote a set of segments of the line AQ such that g_i contains A, g_1 does not contain p_1, g_i contains the end-points of g_{i+1}, and A is the only point common to the g_i.† Let P_{21} denote a point in g_2 such that $AP_{21}p_1$. Let P_{21} correspond to 1/4. For each positive integer $n > 1$ let $c_{n1}, c_{n2}, \cdots, c_{n2^{n-2}}$ denote the positive rational fractions less than $\frac{1}{2}$ which in their lowest terms have the denominator 2^n, and let the second subscripts be chosen so that $c_{ni} < c_{ni+1}$ $(0 < i < 2^{n-2})$. If points $P_{n1}, P_{n2}, \cdots, P_{n2^{n-2}}$ have been put into correspondence with these fractions (P_{ni} corresponding to c_{ni}) by a process previously described, let H_{ni} denote a set of segments $h_1, h_2, h_3, \cdots$ of the line AQ such that h_i contains P_{ni}, h_1 contains no point selected to correspond to a positive fraction less than or equal to $\frac{1}{2}$ with the denominator 2^r $(0 < r < n+1)$ except P_{ni}, h_i contains the end-points of h_{i+1}, and P_{ni} is the only point common to the h_i. Let the points which have been put into correspondence with fractions which

* If A, B, C are points, "ABC" used as a statement means A, B, C are in the order ABC.

† That such a set of segments exists is a consequence of Theorem 1.

have for their denominators powers of 2 not greater than the n-th be denoted by $Q_{n1}, Q_{n2}, \cdots, Q_{n2^{n-1}}$ in such a way that $AQ_{ni}Q_{ni+1}$ $(0 < i < 2^{n-1})$. For each i $(0 < i < 2^{n-1}+1)$, let $P_{(n+1)i}$ denote a point such that $Q_{ni-1}P_{(n+1)i}Q_{ni}$ $(Q_{n0}=A)$ and such that $P_{(n+1)i}$ is in h'_{n+1} where h'_{n+1} denotes g_{n+1} if $i=1$ and h'_{n+1} denotes the $n+1$-st segment H_{jk} if $Q_{ni-1}=P_{jk}$ for some j and some k $(1 < j < n+1; 0 < k < 2^{j-2}+1)$. Let $P_{(n+1)i}$ correspond to $c_{(n+1)i}$. Let $[P]$ denote the set of all the points obtained by continuing this process for $n=3, 4, 5, \cdots$, together with p_1 and P_{21}. Let p_i denote the point of $[P]$ which corresponds to c_i $(i=1, 2, 3, \cdots)$.*

For each point x of AQ † let $f(A,x)=f(x,A)$ denote a number determined as follows:

(1) $f(A,x)=0$ if $x=A$.

(2) If $x \neq A$, let $[p]_x$ denote the set of all points p of the p_i such that Apx, and let $[c]_x$ denote the set of all fractions c such that a point of $[p]_x$ corresponds to c; let $f(A,x)$ denote the upper limit of $[c]_x$.

Then if AxQ, $f(A,x)>0$, and if $Axx'Q$, $f(A,x')-f(A,x) \geqq 0$; indeed, if both x and x' belong to $[P]$, this difference is positive. Also, for each positive integer k there exist points $P_1, P_2, \cdots, P_{2^{k-1}}$, all of which belong to $[P]$, such that $f(A,P_i)=i(\frac{1}{2})^k$.

If V and W are points such that AVC and $W=A$, $W=V$, or AWV, let $f(V,W)=f(W,V)$ be defined as follows: let $T_{DW}(V)$ denote the projection through D of V on QW, and let $T_{CW}(V)$ denote the projection through C of $T_{DW}(V)$ on AQ. Let $f(V,W)=f[A, T_{CW}(V)]$. Then $f(V,W) \geqq 0$, and is zero if and only if $W=V$. If Z is also a point of the segment AC, and if WVZ, then $f(V,W) \leqq f(W,Z) \leqq \frac{1}{2}$. For each positive number e there exists a subinterval s of AC containing W such that W is not a limit point of $AC-s$ and such that if z is a point of s then $f(W,z)<e$. Also, if s denotes a subinterval of AC which contains W and such that W is not a limit point of $AC-s$, then there exists a positive number d such that if y is any point of $AC-C-s$, then $f(W,y)>d$.

If V and W are points of AB, let $\phi(V,W)=\phi(W,V)$ denote the upper

---

* That A is a limit point of the p_i is a consequence of the fact that it follows from the definition of limit point that *a point A of a line l is a limit point of a subset L of l if and only if every segment of l which contains A contains a point of L distinct from A*, and the p_i have been so chosen that every segment of the line AQ which contains A contains one of the p_i.

† If A and Q are distinct points, the Symbol "AQ" unmodified by a word such as "line," "segment," "ray," etc. means the interval AQ; that is, A and Q together with all points X such that AXQ.

limit of $|f(Z,W)-f(Z,V)|$ for all points Z of $AC-C$.* The function ϕ is a distance for points of AB; that is it satisfies the following conditions:

(1) If V and W are points of AB $\phi(V,W)>0$ if $V\neq W$, and $\phi(V,W)=0$ if $V=W$. For if $V\neq W$ $|f(Z,W)-f(Z,V)|>0$ for $Z=V$, and if $V=W$, $f(Z,W)=f(Z,V)$ for each Z.

(2) If V, Y, W are points of AB, then $\phi(V,W)+\phi(W,Y)\geqq\phi(V,Y)$, for $f(Z,V)-f(Z,Y)=f(Z,V)-f(Z,W)+f(Z,W)-f(Z,Y)$, and hence $|f(Z,V)-f(Z,Y)|\leqq|f(Z,V)-f(Z,W)|+|f(Z,W)-f(Z,Y)|$.

(3) If V is a point of AB and is not a limit point of a subset M of AB, then there exists a positive number e such that if z is any point in M, then $\phi(V,z)>e$. For there exists a subinterval s of AB which contains V and which contains no point of M distinct from V and such that V is not a limit point of $AB-s$; hence there exists a positive number e such that if z is in M, then $f(V,z)>e$. By letting $Z=V$, it is seen that $\phi(V,z)$ is at least as great as $f(V,z)$.

(4) If V is a point of AB and is a limit point of a subset M of AB, then for each positive number e there exists a point x of M such that x is distinct from V and $\phi(V,x)<e$.

To establish (4) it will be assumed that AVB. The modifications of the argument necessary to establish the property for $V=A$ and $V=B$ will be pointed out at its conclusion. Let k denote a positive integer such that $(\frac{1}{2})^k<e/4$. Let $P_1, P_2, \cdots, P_{2^{k-1}}$ denote points of $[P]$ such that $f(A,P_i)=i(\frac{1}{2})^k$, then AP_iP_{i+1} $(1\leqq i<2^{k-1})$. Let $F_{CV}(P_i)$ denote the projection through C of P_i on QV, and let x_i denote the projection through D of $F_{CV}(P_i)$ on VC. Then $Vx_1x_2\cdots x_{2^{k-1}}C$. If x is a point of Vx_1 and Z is also in Vx_1, then $f(Z,V)\leqq(\frac{1}{2})^k$ and $f(Z,x)\leqq(\frac{1}{2})^k$, and hence $|f(Z,V)-f(Z,x)|\leqq(\frac{1}{2})^k$. Let b_i denote the intersection of CP_{i-1} $(i=1,2,\cdots,2^{k-1})$ with $DF_{CV}(P_i)$, where $P_0=A$. On account of the order of the P_i and the order DCQ, and since the P_i are all on the A-side of the line DQ, the common part of the interiors of angles x_iDx_{i+1} and b_iQV is interior to angle $P_{i-1}CP_{i+1}$ $(1\leqq i\leqq 2^{k-1}-1)$. Let b_i' denote the projection through Q of b_i on VC. Then for $1\leqq i\leqq 2^{k-1}-1$ if x is a point of Vb_i' and Z is in x_ix_{i+1}, VxZ (or $Z=x$), $T_{CV}(Z)$ is in P_iP_{i+1}, and $T_{Cx}(Z)$ is in $P_{i-1}P_{i+1}$. Hence $(i-1)(\frac{1}{2})^k\leqq f(Z,x)\leqq(i+1)(\frac{1}{2})^k$ and $i(\frac{1}{2})^k$

---

* Cf. E. W. Chittenden's definition of écart in terms of a *Hahn function* which he defines in terms of voisinage in his paper "On the Equivalence of Écart and Voisinage," *Transactions of the American Mathematical Society*, Vol. 18 (1917), pp. 161-166.

$\leqq f(Z, V) \leqq (i+1)(\frac{1}{2})^k$, and hence $| f(Z, V) - f(Z, x) | \leqq 2(\frac{1}{2})^k$. If Z is in $x_{2^{k-1}}C - C$ and x is in $Vb'_{2^{k-1}}$, then $T_{CV}(Z)$ is in $P_{2^{k-1}}Q$ and $T_{Cx}(Z)$ is in $P_{2^{k-1}-1}Q$, and hence $(2^{k-1}-1)(\frac{1}{2})^k \leqq f(Z, x) \leqq \frac{1}{2}$ and $f(Z, V) = \frac{1}{2}$. Then if X_e denotes a point such that VX_eb_i' $(i = 1, 2, \cdots, 2^{k-1})$, $| f(Z, V) - f(Z, x) | \leqq 2(\frac{1}{2})^k$ if Z is in $VC - C$ and x is in VX_e.

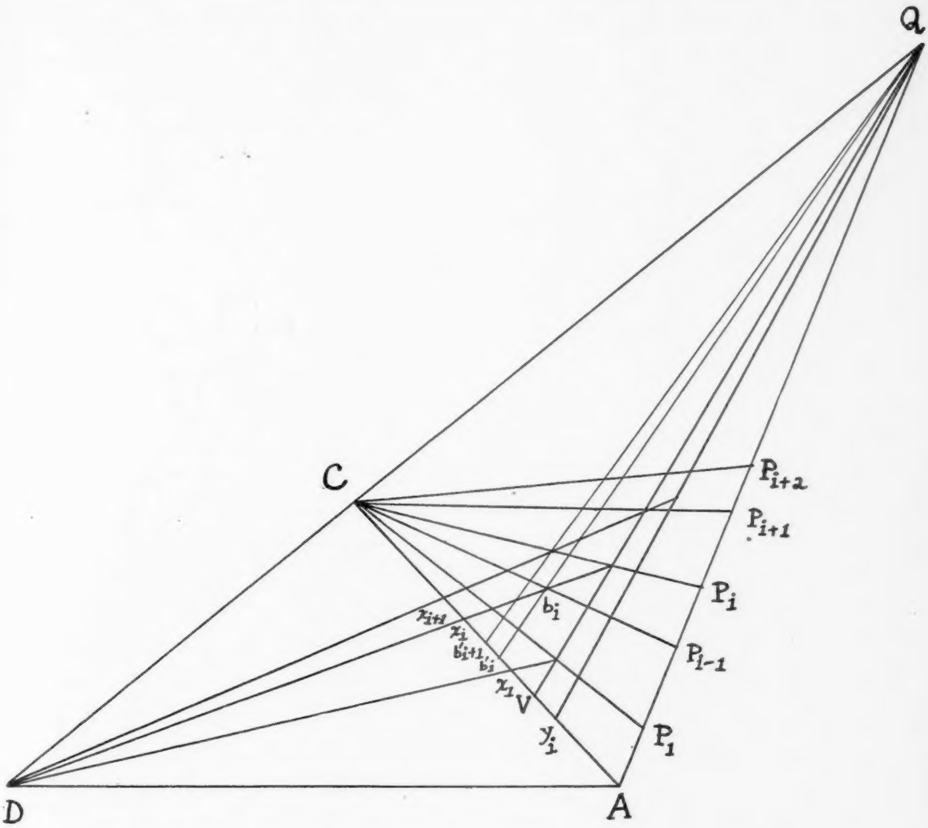

FIG. 1

(for the determination of X_e and Y_e).

Let V_0 denote the projection through D of V on AQ. If AV_0P_1 or if $V_0 = P_1$, then if Z is in VA, $f(Z, V) \leqq (\frac{1}{2})^k$, and if x is a point of VX_e' (where X_e' denotes a point in the orders $VX_e'C$ and $DX_e'P_2$), $f(Z, x) \leqq 2(\frac{1}{2})^k$.

If AP_1V_0, X_e' will be selected differently. Let j denote the greatest integer for which AP_jV_0, and let $a_1, a_2, \cdots, a_j$ denote the points on VA

such that QQ_ia_i, where Q_i is the intersection of VV_0 with CP_i. Let Q'_{i+1} denote the intersection of Qa_i with CP_{i+1} ($P_{j+1}=Q$ if $j=2^{k-1}$). Let s_i denote the interval $a_{i-1}a_i$ ($i=1, 2, \cdots, j+1$), where $a_0=V$ and $a_{j+1}=A$. Let x'_{i+1} denote a point of VC such that $Dx'_{i+1}Q'_{i+1}$. Then if Z is in s_i and Vxx'_{i+1} ($x'_{j+1+1}=C$ if $j\geqq 2^{k-1}-1$), $(i-1)(\tfrac{1}{2})^k\leqq f(Z,V)\leqq i(\tfrac{1}{2})^k$ and

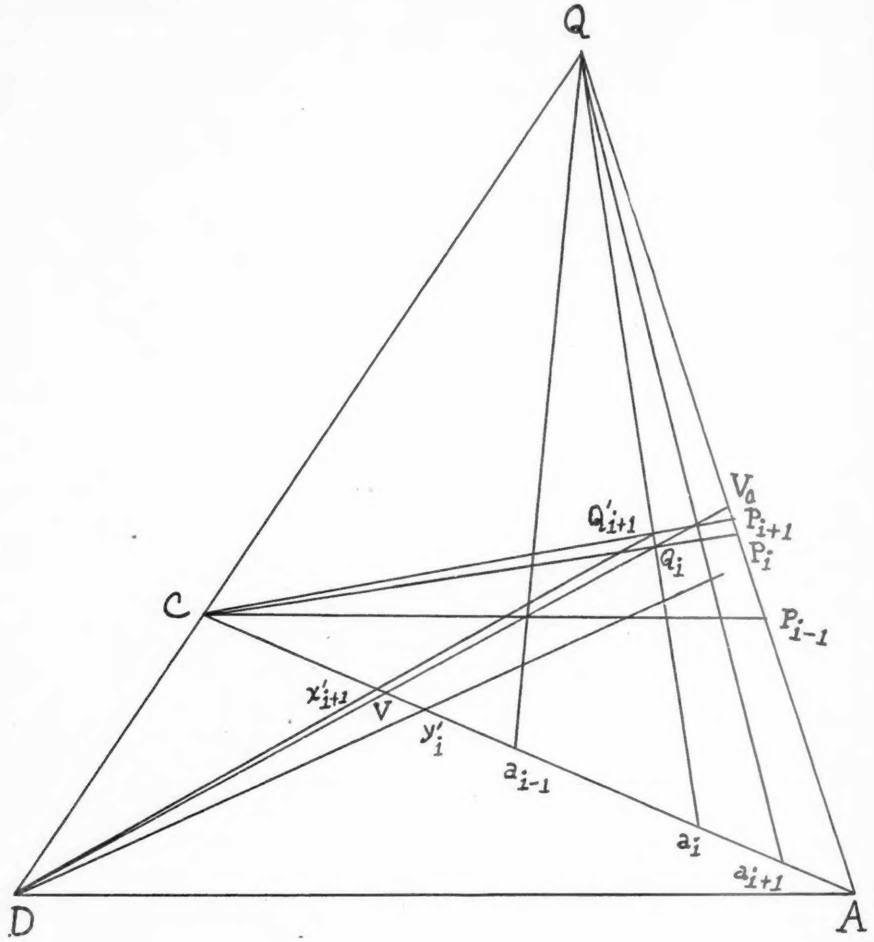

FIG. 2

(for the determination of X_e' and Y_e').

$(i-1)(\tfrac{1}{2})^k\leqq f(Z,x)\leqq(i+1)(\tfrac{1}{2})^k$. Let X_e' denote a point such that $VX_e'x'_{i+1}$ for $i=1, 2, \cdots, j+1$. Then $|f(Z,V)-f(Z,x)|\leqq 2(\tfrac{1}{2})^k$ if VxX_e' and Z is in AV.

Let X_e^* denote a point such that $VX_e^*X_e'$ and $VX_e^*X_e$. Then if x is in VX_e^*, $\phi(V,x) \leqq 2(\frac{1}{2})^k < e$.

In the determination of X_e a set of points $x_1, x_2, \cdots, x_{2^{k-1}}$ was defined. Let y_0 denote a point such that Vy_0A and the segment Qy_0 intersects the ray Dx_1 within the angle P_1CP_2. For $i = 1, 2, \cdots, 2^{k-1} - 1$ let y_i denote a point in the order Vy_iA and such that the segment Qy_i intersects the ray Dx_{i+1} within the angle $P_{i+1}CP_{i+2}$ $(P_{2^{k-1}+1} = Q)$. Let $y_{2^{k-1}} = A$. Let Y_e denote a point such that VY_ey_i $(i = 0, 1, 2, \cdots, 2^{k-1})$. Then if Z is $VC - C$ and x is in VY_e, $| f(Z,V) - f(Z,x) | \leqq 2(\frac{1}{2})^k$, for if Z is in x_ix_{i+1} $(x_0 = V$ and $x_{2^{k-1}+1} = C)$, and $Z \neq C$, and x is in Vy_i, then $i(\frac{1}{2})^k \leqq f(Z,V) \leqq (i+1)(\frac{1}{2})^k$ and $i(\frac{1}{2})^k \leqq f(Z,x) \leqq (i+2)(\frac{1}{2})^k$.

Employing the notation used in the determination of X_e', suppose that V_0 is in AP_3; then if both Z and x are in AV, $| f(Z,V) - f(Z,x | \leqq 3(\frac{1}{2})^k$, for neither $f(Z,V)$ nor $f(Z,x)$ is greater than $3(\frac{1}{2})^k$. Suppose, secondly, that V_0 is not in AP_3, then $j > 2$. For $2 \leqq i \leqq j$, let y_i' denote a point of the segment AV such that the common part of the interiors of angles VDy_i' and a_iQa_{i+1} is within the angle $P_{i-1}CP_{i+1}$. If both x and Z are in Va_2, then neither $f(Z,V)$ nor $f(Z,x)$ is greater than $2(\frac{1}{2})^k$. If x is in Vy_i' and Z is in a_ia_{i+1} $(2 \leqq i \leqq j)$, then neither $f(Z,V)$ nor $f(Z,x)$ is outside the limits $(i-1)(\frac{1}{2})^k$ and $(i+1)(\frac{1}{2})^k$. If V_0 is in AP_3, let $Y_e' = A$. If V_0 is not in AP_3, let Y_e' denote a point such that $VY_e'a_2$ and $VY_e'y_i'$ for each i $(2 \leqq i \leqq j)$. Then if Z is in AV and x is in VY_e', $| f(Z,V) - f(Z,x) | \leqq 3(\frac{1}{2})^k$.

Let Y_e^* denote a point such that $VY_e^*Y_e$ and $VY_e^*Y_e'$, then if x is in VY_e^*, $\phi(V,x) \leqq 3(\frac{1}{2})^k < e$. Let s denote the segment $X_e^*Y_e^*$, then s contains V and a point of M distinct from V, and if x is a point of s, $\phi(V,x) < e$.

If $V = A$, only the point X_e need be determined to establish the property (4) of ϕ; if $V = B$, only the point Y_e' is required.

The distance ϕ for AB will now be used to define a distance for all of S. Let B_0 denote a point such that DBB_0 and AB_0Q, and let R_t and R_v denote points such that $AR_tR_vB_0$. Let L_1 denote the point set consisting of the triangle ABD and its interior. If a is a point of L_1, let t_a denote the projection through R_t of a on AB, and let v_a denote the projection through R_v of a on AB. If b is also a point of L_1, let $d_1(a,b) = d_1(b,a) = \phi(t_a, t_b) + \phi(v_a, v_b)$. Then d_1 is a distance for L_1. Let K_x, K_y, K_z denote three non-collinear points within the triangle ABD, and let L_2 denote the point set consisting of the triangle ABD and its exterior. If p is a point of L_2, let w_p denote the intersection of the ray K_wp $(w = x, y, z)$ with the triangle ABD. If q is also a point of L_2, let

$$d_2(p,q) = d_2(q,p) = (1/3)[d_1(x_p, x_q) + d_1(y_p, y_q) + d_1(z_p, z_q)].$$

Then d_2 is a distance for L_2, and, for points of the triangle ABD, d_2 is identical with d_1. If, now, p and q denote any two points of S, let $d(p,q) = d(q,p)$ be defined as follows:

(1) If both points belong to L_1, $d(p,q) = d_1(p,q)$.

(2) If both points belong to L_2, $d(p,q) = d_2(p,q)$.

(3) If p is in the interior of the triangle ABD and q is in its exterior, let $d(p,q)$ denote the lower limit of $[d_1(p,x) + d_2(x,q)]$ for all points x of the triangle ABD. Then d is a distance for S.

Part II.

Let S^* denote a descriptive plane which contains two points A_0 and B_0 such that the segment A_0B_0 is separable. The assumption that a segment of S^* is separable is not stronger than the assumption that the interior of some triangle of S^* is separable, for it is easily seen, by means of projections, that if the interior of a triangle of a descriptive plane is separable, then each of its sides is separable.

Theorem 3. *Every line* † *of S^* is separable.*

Proof. Let C denote a point such that A_0, B_0, C are non-collinear. Let D and E denote points in the orders CDA_0 and CB_0E, respectively. Let H denote a countable point set everywhere dense on A_0B_0. Let T denote the set of all points X of the ray B_0E such that X is collinear with D and a point of H. The ray B_0E is a subset of $\bar{T}$. Hence it is separable. Similarly the ray B_0C is separable, and therefore so is every segment of it. In view of the argument just given it follows that every line intersecting the line B_0C is separable. If l is a line not intersecting the line B_0C, there exists a line intersecting both l and B_0C and since this line is separable so is l.

Corollary. *If A and B are two distinct points, there exists, on the ray AB, a countable set of points $P_1, P_2, P_3, \cdots$ such that AP_nP_{n+1} and such that if x is any point of the ray AB then either $x = A$ or AxP_n for some n.*

Theorem 4. *If A and B are two distinct points and a and b are two distinct real numbers such that $a < b$, there exists between the points of AB and a subset of the numbers of the interval (a,b) a one-to-one reciprocal*

† In this paper, the word " line," unmodified, is used in the sense of " straight line."

correspondence in which every rational number of (a, b) *corresponds to some point of* AB, A *corresponds to* a, B *corresponds to* b, *and if* c *and* f *are numbers of* (a, b) *corresponding to points* p *and* q, *respectively, of* AB *and* $a < c < f < b$, *then* $ApqB$.

Proof. Let $p_1, p_2, p_3, \cdots$ denote a countable subset of the segment AB everywhere dense on AB and such that $p_i \neq p_j$ if $i \neq j$. Let $c_1, c_2, c_3, \cdots$ denote the set of all the rational numbers of (a, b) distinct from a and from b, $c_i \neq c_j$ if $i \neq j$. Let p_1 correspond to c_1. Let p_{21} and p_{22} denote the p_i with smallest subscripts such that $Ap_{21}p_1p_{22}B$, and let c_{21} and c_{22} denote the c_i with smallest subscripts such that $a < c_{21} < c_1 < c_{22} < b$. Let p_{21} correspond to c_{21} and p_{22} to c_{22}. If j points have been put into correspondence with j numbers in this way, let the points be denoted by $p_{j1}, p_{j2}, \cdots, p_{jj}$ so that $Ap_{ji}p_{ji+1}p_{ji+2}B$ $(1 \leqq i \leqq j-2)$, and let the corresponding c_i be denoted by $c_{j1}, c_{j2}, \cdots, c_{jj}$, using the same subscripts for corresponding points and numbers. Let $k = j + 1$ additional points $p_{k1}, p_{k2}, \cdots, p_{kk}$ and k additional numbers $c_{k1}, c_{k2}, \cdots, c_{kk}$ be selected as follows: p_{k1} is the p_i with smallest subscript such that $Ap_{k1}p_{j1}$; p_{kk} is the p_i with smallest subscript such that $p_{jj}p_{kk}B$, and, for $1 < n < k$, p_{kn} is the p_i with smallest subscript such that $p_{jn-1}p_{kn}p_{jn}$; c_{k1} is the c_i with smallest subscript such that $a < c_{k1} < c_{j1}$; c_{kk} is the c_i with smallest subscript such that $c_{jj} < c_{kk} < b$, and, for $1 < n < k$, c_{kn} is the c_i with smallest subscript such that $c_{jn-1} < c_{kn} < c_{jn}$. Let p_{kn} correspond to c_{kn} $(1 \leqq n \leqq k)$. Let A correspond to a and B to b. If q is a point in the order AqB and q is not one of the p_i, let q correspond to the number which is the lower limit of the rational numbers corresponding to the p_i which are in qB.

Definition. If A, B *are distinct points and* C, D *are distinct points, the term "a segment from* AB *to* CD*" means a segment which has one of its end-points on the segment* AB *and the other on the segment* CD *and which has no other point in common either with* AB *or with* CD.

Definition. If A, B, C *are three non-collinear points, a set* G *of segments from* AB *to* BC *is said to fill the interior of the triangle* ABC *if no two segments of* G *have a point (or an end-point) in common and every point of the interior of the triangle* ABC *belongs to some segment of* G.

Definition. If A, B, C *are three non-collinear points and* D *and* E *are points in the orders* ADB *and* BEC, *respectively, a set* H *of segments from* AD *to* CE *is said to fill the interior of the quadrilateral* $ADEC$ *if no two*

segments of H have a point (or an end-point) in common and every point of the interior of the quadrilateral belongs to some segment of H.

In order to show the existence of a set G of segments from AB to BC filling the interior of the triangle ABC, let F denote a point such that ACF, and let G consist of all the segments from AB to BC which are subsets of lines through F. In order to show the existence of a set H of segments from AD to CE filling the interior of the quadrilateral $ADEC$, let K denote a point in the order EDK † and let H denote the set of all segments h from AD to CE such that h is a subset of a ray which either starts at F and is interior to the angle KFA or starts at K and is interior to the angle FKE or coincides with the ray KF.

The definitions just given will be of use in proving the following theorem:

THEOREM 5. *If O is a point of S^*, there exists between the points of S^* and an everywhere dense subset of the number-plane a one-to-one continuous correspondence in which the image of each line of S^* which contains O is an everywhere dense subset of a line of the number plane through the origin, and the image of each line of S^* is an everywhere dense subset of an open curve of the number-plane.*‡

Proof. Let A_1 and B_1 denote two distinct points each distinct from O and such that O, A_1, B_1 are non-collinear. Let C_1 and D_1 denote points in the orders A_1OC_1 and B_1OD_1, respectively. Let $A_2, A_3, A_4, \cdots$; $B_2, B_3, B_4, \cdots$; $C_2, C_3, C_4, \cdots$; $D_2, D_3, D_4, \cdots$ denote sequences of points on the rays OA_1, OB_1, OC_1, OD_1, respectively, and such that each of the sequences $A_1, A_2, A_3, A_4, \cdots$; $B_1, B_2, B_3, B_4, \cdots$; $C_1, C_2, C_3, C_4, \cdots$; $D_1, D_2, D_3, D_4, \cdots$ has, with respect to the corresponding ray, the properties stated in the Corollary to Theorem 3. Let the points of OA_1 correspond to numbers of the interval $(0, 1)$ in the manner indicated in Theorem 4 and so that A_1 corresponds to 1. For each positive integer i let the points of A_iA_{i+1} correspond to numbers of the interval $(i, i+1)$ in the same manner and so that A_i corresponds to i. Let the points of A_1B_1 correspond in this way to numbers of the interval $(0, 1)$ so that A_1 corresponds to 0, and let the points of B_1C_1 correspond in the same manner to numbers of the interval $(1, 2)$, B_1 corresponding to 1. If p is a point of C_1D_1, and q is a point of A_1B_1 collinear with O and p, let x

† In addition, let K be on the D-side of the line AC, and let F be in the order ACF and within the angle ADE.

‡ Cf. R. L. Moore, "On a Set of Postulates Which Suffice to Define a Number Plane," *Transactions of the American Mathematical Society*, Vol. 16 (1915), pp. 27-32.

denote the number to which q corresponds ($x = 0$ if $q = A_1$), and let p correspond to $2 + x$. If p is a point of $D_1A_1 - A_1$, and q is a point of B_1C_1 collinear with O and p, let x denote the number to which q corresponds, and let p correspond to $2 + x$. In this way, the points of C_1D_1 correspond to numbers of the interval $(2, 3)$, and the points of $D_1A_1 - A_1$ to numbers of $(3, 4)$, D_1 corresponding to 3.

For each rational number u $(0 \leqq u < 4)$ let r_u denote the ray Op_u where p_u is the point on the quadrilateral $A_1B_1C_1D_1$ which corresponds to u. For each u let $p_{u1}, p_{u2}, p_{u3}, \cdots$ denote a sequence of points on the ray r_u such that $Op_{ui}p_{ui+1}$ and such that any point of the ray r_u except O is between O and some p_{ui}. Let $p_{0i} = A_i$, $p_{1i} = B_i$, $p_{2i} = C_i$, $p_{3i} = D_i$.

Let the interiors of the triangles A_1OB_1, B_1OC_1, C_1OD_1 be filled with sets G_1, G_2, G_3, of segments from OA_1 to OB_1, from OB_1 to OC_1, from OC_1 to OD_1, respectively. For each positive number x to which there corresponds a point of the ray OA_1 let q_x denote that point. For $x < 1$, let g_{1x} denote the segment of G_1 which has q_x for one end-point, let g_{2x} denote the segment of G_2 which has an end-point in common with g_{1x}, and let g_{3x} denote the segment of G_3 which has an end-point in common with g_{2x}. Let $g'_{4,1/2}$ denote the segment from the end-point of $g_{3,1/2}$ on OD_1 to $q_{1/2}$ (this segment is interior to the triangle D_1OA_1). Let q_{jx} $(1 \leqq j < 4)$ denote the end-point of g_{jx} on Op_j $(p_1 = B_1, p_2 = C_1, p_3 = D_1)$.

For each positive integer n, by definition, $p_{(4-1/n)}$ denotes the point on D_1A_1 which corresponds to $4 - 1/n$. Let $q_{(4-1/2),1/2}$ denote the intersection of $g'_{4,1/2}$ with $r_{(4-1/2)}$. Let G_4 denote a set of segments from OD_1 to $Op_{(4-1/2)}$ which includes the segment $q_{3,1/2}q_{(4-1/2),1/2}$ and which fills the interior of the triangle $D_1Op_{(4-1/2)}$. That G_4 may contain the segment specified and fill the interior of the triangle follows from the fact that it may consist of a set of segments filling the interior of the triangle $q_{3,1/2}Oq_{(4-1/2),1/2}$ plus the segment $q_{3,1/2}q_{(4-1/2),1/2}$ plus a set of segments filling the interior of the quadrilateral $D_1q_{3,1/2}q_{(4-1/2),1/2}p_{(4-1/2)}$.

Let g_{4x} denote the segment of G_4 which has an end-point in common with g_{3x}. Let $q_{(4-1/2),x}$ denote the end-point of g_{4x} on $Op_{(4-1/2)}$. Let $g'_{5,1/3}$ and $g'_{5,2/3}$ denote the segments $q_{(4-1/2),1/3}q_{1/3}$ and $q_{(4-1/2),2/3}q_{2/3}$, respectively, and let $g'_{5,1/2}$ denote the segment $q_{(4-1/2),1/2}q_{1/2}$. Let $q_{(4-1/3),1/3}$, $q_{(4-1/3),1/2}$, $q_{(4-1/3),2/3}$ denote the intersections of $Op_{(4-1/3)}$ with $g'_{5,1/3}$, $g'_{5,1/2}$, $g'_{5,2/3}$, respectively. Let G_5 denote a set of segments from $Op_{(4-1/2)}$ to $Op_{(4-1/3)}$ which fills the interior of the triangle $p_{(4-1/2)}Op_{(4-1/3)}$ and which includes the subsegments of $g'_{5,1/3}$, $g'_{5,1/2}$, $g'_{5,2/3}$ which are within this triangle.

If for an integer n and for each integer m $(1 < m < n)$ a set G_{2+m} of segments from $Op_{[4-1/(m-1)]}$ to $Op_{(4-1/m)}$ filling the interior of the triangle whose vertices are the end-points of the segments just named has been selected, in the manner described for $m = 2$, 3, let $c_{n1}, c_{n2}, \cdots, c_{nN}$ denote the proper rational fractions which in their lowest terms have a denominator $\leqq n$ (N denotes the number of these fractions), and let $g'_{2+n,c_{nj}}$ denote the

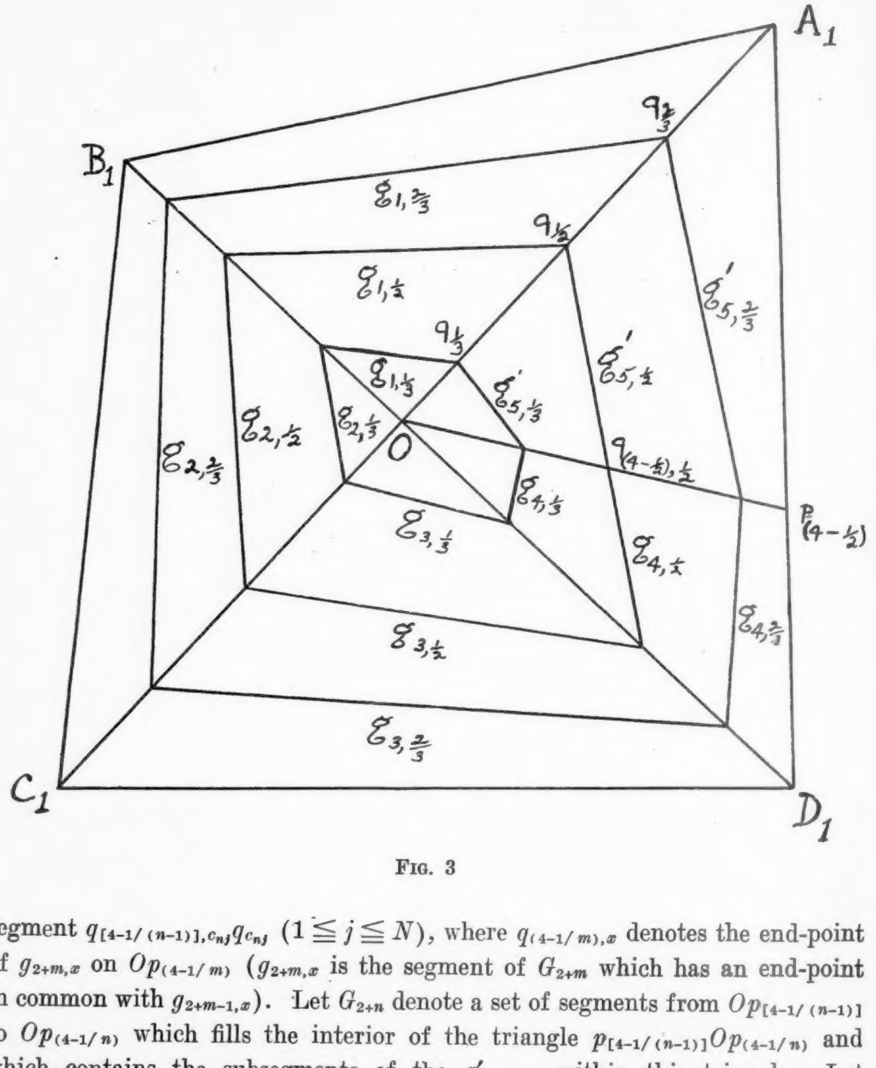

FIG. 3

segment $q_{[4-1/(n-1)],c_{nj}}q_{c_{nj}}$ $(1 \leqq j \leqq N)$, where $q_{(4-1/m),x}$ denotes the end-point of $g_{2+m,x}$ on $Op_{(4-1/m)}$ ($g_{2+m,x}$ is the segment of G_{2+m} which has an end-point in common with $g_{2+m-1,x}$). Let G_{2+n} denote a set of segments from $Op_{[4-1/(n-1)]}$ to $Op_{(4-1/n)}$ which fills the interior of the triangle $p_{[4-1/(n-1)]}Op_{(4-1/n)}$ and which contains the subsegments of the $g'_{2+n,c_{nj}}$ within this triangle. Let $g_{2+n,x}$ denote the segment of G_{2+n} which has an end-point in common with $g_{2+n-1,x}$.

If c is a proper rational fraction which in its lowest terms has the denominator m, the point set $g_{1c} + g_{2c} + g_{3c} + \cdots + g'_{2+m,c}$ plus the end-points of these segments is a polygon which has q_c as a vertex or as an inner point of one of its sides, for q_c is a common end-point of g_{1c} and $g'_{2+m,c}$, and in the sets of segments described in the preceding paragraph g_{kc} for $k \geqq 2 + m$ is a subsegment of $g'_{2+m,c}$. For each real number x to which there corresponds a point of the segment OA_1, let I_{ix} denote the interval obtained from g_{ix} by adding to the latter its end-points. For each such number x, let $K_x = \sum_{i=1}^{i=\infty} I_{ix}$. Then q_x is a limit point of the common part of K_x and the interior of the triangle D_1OA_1, and no other point of the ray OA_1 is a limit point of K_x. The subset of K_x within the triangle D_1OA_1 will be called a broken line from OD_1 to OA_1 and q_x and q_{3x} will be called its end-points; the collection of all such broken lines—that is, for all values of x corresponding to points of OA_1—will be said to fill the interior of the triangle D_1OA_1.

For each integer t $(t > 1)$ let T denote the number of rational numbers greater than or equal to 0 and less than 4 which are integers or which in their lowest terms have a denominator less than or equal to t. Let these numbers be denoted by $b_{t1}, b_{t2}, \cdots, b_{tT}$ in such a way that $b_{ti} < b_{ti+1}$ $(0 < i < T)$. Then $b_{t1} = 0$. Let M_1 denote the quadrilateral $A_1B_1C_1D_1$ and let M_t denote the polygon with the vertices $P_{t1}, P_{t2}, \cdots, P_{tT}$ which are points determined as follows: on the ray $r_{b_{ti}}$ there has been selected a sequence of points $p_{b_{ti}j}$ $(j = 1, 2, 3, \cdots)$ such that $Op_{b_{ti}j}p_{b_{ti}j+1}$ and such that any point of the ray except O is between O and one of these p's. Let P_{t_i} denote the first of the $p_{b_{ti}j}$ outside of M_{t-1}. Then $P_{t1} = A_t$.

It is to be noted that M_t incloses M_{t-1}. It may be seen that the rays from O through the vertices of M_t intersect M_{t-1} in such a way that if Q_{ti} denote the point on M_{t-1} such that $OQ_{ti}P_{ti}$, then the interval $Q_{ti}Q_{ti+1}$ $(Q_{tT+1} = Q_{t1})$ is a subinterval of a side of M_{t-1}. Let V_{ti} denote the quadrilateral with vertices $P_{ti}, P_{ti+1}, Q_{ti+1}, Q_{ti}$ $(P_{tT+1} = P_{t1})$. The interiors of the V_{ti} are mutually exclusive. Let the interior of V_{ti} $(0 < i < T)$ be filled with a set of segments from $Q_{ti}P_{ti}$ to $Q_{ti+1}P_{ti+1}$. Let the interior of the quadrilateral V_{tT} be filled with a set of broken lines from $Q_{tT}P_{tT}$ to $Q_{t1}P_{t1}$, in a manner analogous to that in which the interior of the triangle D_1OA_1 was filled with broken lines from OD_1 to OA_1; that is, with a set of broken lines each consisting of a finite or infinite number of segments such that each point of the segment $Q_{tT}P_{tT}$ and each point of the segment $Q_{t1}P_{t1}$ is an end-point of one and only one broken line of the set, each point of the in-

terior of V_{tT} belongs to one and only one element of the set, and such that the set has the property now to be defined: for each real number x $(t - 1 < x < t)$ to which there corresponds a point of the segment $A_{t-1}A_t$, let I_{t1x} denote the interval obtained by adding its end-points to the segment of the set filling the interior of V_{t1} which has q_x for one end-point, for $1 < i < T$ let I_{tix} denote the interval obtained by adding its end-points to the segment of the set filling the interior of V_{ti} which has an end-point in common with $I_{t,i-1,x}$; let I_{tTx} denote the point set obtained by adding its end-points to the broken line of the set filling the interior of V_{tT} which has an end-point in common with $I_{t,T-1,x}$, then q_x belongs to I_{tTx}. If $0 < t - 1 < x < t$, let K_x denote the point set defined as follows: $K_x = \sum_{i=1}^{i=T} I_{tix}$. For each positive integer x, let K_x denote the polygon M_x.

Then if p is any point of S^* except O, there is one and only one positive number x such that p belongs to K_x; for each p let x_p denote this number. The point in which the ray Op intersects the quadrilateral $A_1B_1C_1D_1$ corresponds to a number, positive or zero and less than 4, by virtue of the correspondence between the points of this quadrilateral and a subset of the numbers of the interval $(0, 4)$; let a_p denote this number. Let p correspond to that point of the number-plane which has the polar coördinates (r_p, ϕ_p) where $r_p = x_p$ and $\phi_p = \frac{1}{2}\pi a_p$, and let O correspond to the pole.

That the subset of the number-plane which corresponds to the points of S^* is everywhere dense in the number-plane is evident from the fact that it contains all the points whose polar coördinates are such that r is rational and ϕ is a rational multiple of $\pi/2$. It can be seen from the manner in which the numbers x_p and a_p are determined by a point p of S^* that p is a sequential limit point of a sequence of distinct points $p_1, p_2, p_3, \cdots$ if and only if $\lim_{i=\infty} x_{p_i} = x_p$ and $\lim_{i=\infty} a_{p_i} = a_p$, unless $p = O$ in which case it is only necessary that $\lim_{i=\infty} x_{p_i} = 0$. Hence the correspondence is continuous. If p is a point of S^* distinct from O, the points of the line Op correspond to an everywhere dense subset of the line in the number-plane whose equation is $\phi = \frac{1}{2}\pi a_p$. Let l denote a line of S^* which does not contain O, and let L denote the image of l in the number plane. It may be noted that if x' is any positive real number, then there exists a number $x > x'$ such that x corresponds to a point of the ray OA_1 and l intersects K_x; hence L is not bounded. Let A and B denote two distinct points of l such that AB contains no point of the ray OA_1. The image of AB in the number-plane can be covered with a simple chain of regions each region of which is the interior of a simple closed curve of diameter less than any previously assigned positive

number d, for AB can be covered with a simple chain of quadrilaterals * whose sides are subintervals of the K_x and of lines through O. These quadrilaterals may be chosen so that their images are of diameter less than d and so that each incloses at least one point of AB, and since the image of each K_x is a subset of the circle $r = x$, the images of these quadrilaterals are subsets of simple closed curves composed of segments of lines through the pole and of circles with center at the pole. If e is a positive number and C' is a chain of the type described covering the image of AB, then there exists another such chain covering the image of AB and such that the boundary of each region of the second chain is a subset of some region of C' and is of diameter less than e. It follows, since $\bar{L}$ contains at most one point of the polar axis, that any two points of $\bar{L}$ belong to a continuum lying in L and hence $\bar{L}$ is a continuum. That the omission of any point of $\bar{L}$ divides it into two mutually separated connected point sets may be seen by observing that the line through the pole and a point of L divides the number-plane into two mutually separated domains each of which contains a connected subset of $\bar{L}$ (the connectedness of the indicated subsets of $\bar{L}$ can be established in the same way in which the connectedness of L was shown). Hence $\bar{L}$ is an open curve. Furthermore, $\bar{L}$ has the property that no line through the pole contains more than one point of it, and if a line whose equation is $\phi = m\pi/n$ (m and n integers) intersects $\bar{L}$ the point of intersection corresponds to a point of l.

Part III.

Definition. If the points of a line l of a descriptive plane are divided into two sets S_1 and S_2 such that no point of one set is between two points of the other and no point P is between every point of S_1 distinct from P and every point of S_2 distinct from P, the pair of point sets S_1 and S_2 will be called a "gap" in l and will be said to divide l into S_1 and S_2.

Definition. If g is a gap in a line l of a descriptive plane which divides l into the point sets S_1 and S_2 and p and q are points of l, pqg will mean that both p and q are in S_1 (S_2) and there exists a point q' in S_2 (S_1) such that pqq'; pgq will mean that p is in one of the sets S_1, S_2 and q is in the other. A triangle will be said to inclose g if it incloses a point of S_1 and a point of S_2.

---

* A finite set of polygons $Z_1, Z_2, \cdots, Z_k$ is said to be a simple chain of polygons if the interior of Z_i has a point in common with the interior of Z_{i+j} ($1 \leqq i < k$; $1 \leqq j \leqq k - i$) if and only if $j = 1$. A set Z of polygons is said to cover a point set M if and only if every point of M is inclosed by some polygon of Z.

In his paper "Concerning a Non-Metrical Pseudo-Archimedean Axiom,"* R. L. Moore employed a pseudo-archimedean axiom which may be briefly stated: *if g is a gap in a line l and A and B are two distinct points on the same side of the line l, then if a triangle incloses g it incloses a point C on the opposite side of l from A and B such that the triangle ABC incloses g.*

Definition. If it is true of a descriptive plane that this pseudo-archimedean axiom is satisfied at no gap in that plane, then the plane is said to be "quasi-complete."

There will now be exhibited an example of a quasi-complete, metric, descriptive plane which is complete and in which it is possible to define congruence of segments in such a way that Moore's set C of congruence axioms † is satisfied, but which is not separable nor locally compact.

The descriptive plane is one described by Hilbert in "Ueber den Satz von der Gleichheit der Basiswinkel im Gleichschenkligen Dreieck" ‡; the definition of congruence, however, is different from the one used by Hilbert. The points of the plane are all the pairs (x, y) of elements of the set T of all the ideal numbers of the form $\alpha = a_0 t^n + a_1 t^{n+1} + \cdots$, where t is a parameter, the number of terms in α is finite or infinite, a_0 $(a_0 \neq 0)$, $a_1, \cdots$ are any real numbers, and n is any integer, positive, negative or zero. An element of T is positive if the corresponding a_0 is positive; $x > y$ if $x - y$ is positive; $y < x$ means $x > y$. To obtain a congruence satisfying Moore's axioms C, let AB be congruent to DE if and only if

$$[(x_A - x_B)^2 + (y_A - y_B)^2]^{1/2} = [(x_D - x_E)^2 + (y_D - y_E)^2]^{1/2}. \S$$

Collineation and order are defined in terms of the elements of T just as they are defined in the number-plane in terms of real numbers. Let H denote this plane.

In order to show that *H is not separable*, it suffices to establish the existence of a non-separable segment in H. For each real number b $(0 < b < 1)$ let s_b denote the segment whose end-points are $(b, 0)$ and $(b + t, 0)$.

---

* *Bulletin of the American Mathematical Society*, Vol. 22 (1916), pp. 225-236.

† See R. L. Moore, "Sets of Metrical Hypotheses for Geometry," *Transactions of the American Mathematical Society*, Vol. 9 (1908), pp. 487-512. Concerning the equivalence, in a descriptive plane, of C and Hilbert's congruence axioms, see Moore's Theorem 1, *ibid.*, § 6, p. 504. That the proposition that every segment has a middle point is a consequence of Moore's order and congruence axioms has been shown by the present author in his paper "Concerning a Set of Metrical Hypotheses for Geometry," *Annals of Mathematics* (2), Vol. 29 (1928), pp. 229-231.

‡ *Grundlagen der Geometrie*, Anhang II, Teubner, 5. Auflage, 1922.

§ T contains $(\alpha)^{1/2}$ if $a_0 > 0$ and n is even.

No two of these segments have a point in common, for if $0 < b < b' < 1$, then $b' > b + t$. Each of these segments is a subset of the segment whose end-points are $(0, 0)$ and $(1, 0)$, and hence the latter segment, since it contains an uncountable set of mutually exclusive segments, is not separable.

If H satisfied the pseudo-archimedean axiom quoted, it would follow from the results of Moore's paper on this axiom that H is separable; hence H does not satisfy this axiom. If it can be shown that all the gaps in H must satisfy this axiom if one does, then it follows that H *is quasi-complete*. A translation of axes disturbs neither collineation nor order in H, and Hilbert has shown the existence of transformations of H into itself which disturb neither collineation nor order and which transform any given line through the origin into any other given line through the origin. Hence, in order to show that H is quasi-complete, it suffices to show that if g is any gap in the x-axis and g_1 is a particular gap in the x-axis, then there exists a transformation of H into itself which disturbs neither order nor collineation and which transforms g into g_1.

Let g denote a gap in the x-axis.* It may be supposed that one of the sets into which g divides the axis consists entirely of points whose abscissas are positive, for, if this is not the case, consider the gap h which divides the axis into the two sets each of which consists of the points whose abscissas are the negatives of the abscissas of the points in one of the two sets into which g divides the axis. Let S_1 and S_2 denote the sets into which g divides the axis, and let S_2 denote that one of these sets which consists entirely of points whose abscissas are positive. Let V denote the subset of T which consists of all the elements v of T such that $(v, 0)$ is in S_1, and let W denote the subset of T consisting of all the elements w of T such that $(w, 0)$ is in S_2. Then every element of T belongs either to V or to W, every element of W is greater than any element of V, and there is neither a greatest element of V nor a least element of W. Hence V contains some positive elements. Let V^* denote the subset of V which consists of all the positive elements of V.

Let N denote the greatest integer—if one exists—such that if $v = a_0 t^n + a_1 t^{n+1} + \cdots$ is in V^* then $n \geqq N$. Such an integer N exists, for suppose, on the contrary, that if m is any integer there exists an element $v = a_0 t^n + a_1 t^{n+1} + \cdots$ of V^* such that $n < m$; since v is in V^* $a_0 > 0$, and it follows that V contains an element greater than any preassigned element of T, contrary to hypothesis. Then V^* contains elements of the form $v = a_0 t^N + a_1 t^{N+1} + \cdots$. Let V_N denote the set of all elements of V^* of this form, then if v is an element of V there exists an element v' of V_N such that

* **Every line of H contains a gap, for H is non-archimedean.**

12

$v' > v$. Let A_i $(i = 0, 1, 2, 3, \cdots)$ denote the set of all real numbers a such that a is the coefficient of the $N + i$-th power of t in some element of V_N. For each positive or zero integer i for which A_i has a finite upper limit, let $\bar{A}_i$ denote that upper limit. There is at least one value of i for which A_i fails to have a finite upper limit, for if all these upper limits are finite, the element $\bar{A}_0 t^N + \bar{A}_1 t^{N+1} + \bar{A}_2 t^{N+2} + \cdots$ of T is either the greatest element of V or the least element of W, contrary to hypothesis. Let k denote the least integer, positive or zero, for which A_k has no finite upper limit. Let $b = 0$ if $k = 0$, and if $k > 0$ let $b = \bar{A}_0 t^N + \bar{A}_1 t^{N+1} + \cdots + \bar{A}_{k-1} t^{N+k-1}$. Suppose that b is in V. Let $v_j = b + jt^{N+k}$ for each positive integer j, and let $z_j = b + (1/j) t^{N+k-1}$; then if v is in V and w is in W there exists an integer j such that $w > z_j > v_j > v$ (v_j is in V and z_j is in W for all j). Suppose, secondly, that b is not in V, then $k > 0$. Let r denote the least positive or zero integer less than k for which $c = \bar{A}_0 t^N + \bar{A}_1 t^{N+1} + \cdots + \bar{A}_r t^{N+r}$ does not belong to V. Let $w_j = c - jt^{N+r+1}$, and let $v_j' = c - (1/j) t^{N+r}$. Then w_j is in W and v_j' is in V, and if v is in V and w is in W there exists a positive integer j such that $w > w_j > v_j' > v$.

Let V_1 denote the subset of T consisting of all elements v of T such that $v < jt^2$ for some positive integer j, and let W_1 denote the subset of T consisting of all elements w of T such that $w > (1/j)t$ for some positive integer j. Then every element of T belongs either to W_1 or to V_1, neither of these two sets contains a greatest element of itself nor a least element of itself, and each element of W_1 is greater than any element of V_1. Let g_1 denote the gap in the x-axis determined by the point sets S^*_1 and S^*_2 consisting of all the points of the x-axis whose abscissas are in V_1 and all the points of the x-axis whose abscissas are in W_1, respectively. A transformation of H into itself which disturbs neither collineation nor order and under which g_1 is the image of g is the transformation under which the image of a point (x, y) is the point (x', y'), where x' and y' are determined as follows:

(1) If b is in V, $x' = (x - b)/t^{N+k-2}$,
and $y' = y/t^{N+k-2}$.

(2) If b is in W, $x' = (c - x)/t^{N+r-1}$,
and $y' = -y/t^{N+r-1}$.

In case (1), S_1 is transformed into S^*_1, and, in case (2), S_1 is transformed into S^*_2.

A complete, compact space is separable.† Since no segment of H is

† F. Hausdorff, *Grundzüge der Mengenlehre* (1914).

separable, H is not locally separable. It follows that if H is complete, then H is not locally compact.

The plane H is complete. If $z = a_0 t^n + a_1 t^{n+1} + \cdots$, let the function $f(z)$ be defined as follows:

$$\begin{aligned} f(z) &= 1, \quad \text{if } n \leqq 1; \\ f(z) &= 1/n, \quad \text{if } n > 1, \end{aligned}$$

and let $f(0) = 0$. If A is a point of H and B is a point of H, let $d(A, B) = d(B, A) = f(x_A - x_B) + f(y_A - y_B)$. Then d is a distance for points of H; that is, it satisfies the three conditions:

(1) $d(A, B) = 0$ if $A = B$, and $d(A, B) > 0$ if $A \neq B$.

(2) $d(A, C) = f(x_A - x_B + x_B - x_C) + f(y_A - y_B + y_B - y_C) \leqq f(x_A - x_B) + f(x_B - x_C) + f(y_A - y_B) + f(y_B - y_C) = d(A, B) + d(B, C)$, for it follows from the definition of $f(z)$ that $f(z + z') \leqq f(z) + f(z')$.

(3) A point A is a sequential limit point of a sequence of distinct points A_i $(i = 1, 2, 3, \cdots)$ if and only if $\lim\limits_{i \to \infty} d(A, A_i) = 0$. Let $A = (x_A, y_A)$ and let $A_i = (x_i, y_i)$. That d satisfies the second part of condition (3) is verified by observing that if A is a sequential limit point of the A_i, then either all but a finite number of the points $(0, y_i)$ are identical with $(0, y_A)$ or $(0, y_A)$ is the sequential limit point of the points $(0, y_i)$; similarly for the point $(x_A, 0)$ and the points $(x_i, 0)$. Hence if e is any positive element of T, there exists a positive integer m such that if $i > m$ then $-e < x_A - x_i \pm (y_A - y_i) < e$, and therefore the least exponent of t in $x_A - x_i$ or in $y_A - y_i$ is not less than the least exponent of t in e. To verify that d satisfies the first part of condition (3), note that if $\lim d(A, A_i) = 0$, then for each positive integer n the rectangle with vertices $(x_A + t^n,\ y_A + t^n)$, $(x_A + t^n,\ y_A - t^n)$, $(x_A - t^n,\ y_A - t^n)$, $(x_A - t^n,\ y_A + t^n)$ incloses all but a finite number of the A_i, and note that if a triangle incloses A it also incloses at least one of these rectangles. To show that H is complete, let $p_1, p_2, p_3, \cdots$ $[p_i = (x_i, y_i)]$ denote a sequence of distinct points such that if h is any positive real number there exists a positive integer m such that if k is any positive integer then $d(p_m, p_{m+k}) < h$, then there exists a point $p_0 = (x_0, y_0)$ which is a sequential limit point of the p_i. For each integer $r > 1$, let m_r denote a positive integer such that $d(p_{m_r}, p_{m_r+k}) < 1/r$ for every positive integer k. Then if $x_{m_r} - x_{m_r+k} = a_0 t^n + a_1 t^{n+1} + \cdots$ and $y_{m_r} - y_{m_r+k} = b_0 t^{n'} + b_1 t^{n'+1} + \cdots$, it follows that $n > r$ and $n' > r$. If $x_{m_2} = 0$ or if the lowest exponent of t in x_{m_2} is greater than 2, let $x_0 = 0 \cdot t$

$+ a_3't^3 + a_4't^4 + \cdots$ (where a_r' is the coefficient of t^r in x_{m_r}). If $x_{m_2} = a_0t^N + a_1t^{N+1} + \cdots$ where $N \leqq 2$, let $x_0 = a_0t^N + a_1t^{N+1} + \cdots + a_2't^2 + a_3't^3 + a_4't^4 + \cdots$ where a_r' is defined as before. Let y_0 denote the element of T similarly determined by the y_{m_r}.

Part IV.

Let T denote an uncountable, well-ordered sequence of parameters t such that no countable subsequence of T runs through T. If t is an element of T, let $t^0 = 1$. A polynomial with real coefficients in a finite number of the elements of T will be said to be equal to 0 if and only if the expression obtained by replacing the different elements of T occurring in it by independent real variables is identically zero. An expression $at_1^{n_1}t_2^{n_2}\cdots t_s^{n_s}$ (the exponents are positive or zero integers) is said to be *lower* than the expression $bt_1^{m_1}t_2^{m_2}\cdots t_s^{m_s}$ $(a \neq 0 \neq b)$ if the first of the differences $n_s - m_s, n_{s-1} - m_{s-1}, \cdots, n_2 - m_2, n_1 - m_2$ which is not zero is negative. Then a non-zero polynomial with real coefficients in a finite number of elements of T may be represented by $a_1Y_1 + a_2Y_2 + \cdots + a_kY_k$ where the Y_i $(i = 1, 2, \cdots, k)$ represent power products of a finite number of elements of T, Y_i is lower than Y_{i+1} $(0 < i < k)$, and the a_i $(a_1 \neq 0)$ are real numbers. Such a polynomial will be said to be less than zero if $a_1 < 0$ and greater than zero if $a_1 > 0$. If P and Q are such polynomials and $P > 0$ and $Q > 0$, then $PQ > 0$ and $P + Q > 0$, for the lowest term in PQ is the product of the lowest term in P by the lowest term in Q, and the lowest term in $P + Q$ is the lowest term in P, the lowest term in Q, or it is $(a_1 + b_1)Y_1$ where a_1Y_1 is the lowest term in P and b_1Y_1 is the lowest term in Q. The statement $P > Q$ means $P - Q > 0$. Then if F is also a polynomial of the same sort as P and Q, it follows from $P > Q$ and $Q > F$ that $P > F$, for if $P - Q > 0$ and $Q - F > 0$, then $P - F = P - Q + Q - F > 0$.

Let R denote the field composed of all rational expressions r such that r contains only a finite number of elements of T, has real coefficients, and does not involve division by zero. Then an element r of R is either a polynomial or it can be expressed as the quotient of two polynomials of R. The relation "greater than" has already been defined for the polynomials of R. If r is an element of R which is not a polynomial, let $r = P/Q$ where P and Q are polynomials of R, then $r > 0$ will mean either $P > 0$ and $Q > 0$ or $P < 0$ and $Q < 0$; $r = 0$ means $P = 0$ and $Q \neq 0$. If r and r' are elements of R, $r > r'$ means $r - r' > 0$. If r, r', r'' are elements of R, there exists

in R a polynomial $Q > 0$ and polynomials P, P', P'' such that $r = P/Q$, $r' = P'/Q$, and $r'' = P''/Q$. Hence if $r > r'$ and $r' > r''$, then $r > r''$. Let $r' < r$ mean $r > r'$. If r is an element of R and r' is an element of R, $r' \neq r$ means $r - r' \neq 0$. In respect of rational operations and relations *greater than* and *less than* the elements of R satisfy all of the rules 1-16 given by Hilbert on pages 35 and 36 of *Grundlagen der Geometrie.* It is this fact which justifies the statement made below that the plane G, there defined, is a descriptive plane.

Let G denote the plane whose points are all the pairs (x, y) of elements of R, in which a line consists of all the points whose coördinates satisfy an equation of the form $Ax + By + C = 0$ (A, B, C are elements of R and not both A and B are zero), and in which three collinear points N, N', N'' are said to be in the order $NN'N''$ when:

(1) Their abscissas x, x', x'', respectively, are such that $x > x' > x''$ or $x < x' < x''$ if $B \neq 0$ in the equation of the line containing the three points.

(2) Their ordinates y, y', y'', respectively, are such that $y > y' > y''$ or $y < y' < y''$ if $B = 0$. Then G is a descriptive plane, but G *is not metric,* for the segment between the points $(0, 0)$ and $(1, 0)$ contains no countable set of points of which $(0, 0)$ is a limit point. For let $[N]$ denote a countable set of points of this segment and let $[t]_N$ denote the set of all elements t of T such that t occurs in the abscissa of some point of $[N]$, then $[t]_N$ is countable, and, by hypothesis, there exists an element t' of T such that t' follows t in T if t is in $[t]_N$. It follows that t', as an element of R, is less than the abscissa of any point of $[N]$, and hence the point $(t', 0)$ is between $(0, 0)$ and every point of $[N]$.

THE UNIVERSITY OF TEXAS.

CONCERNING METRIC COLLECTIONS OF CONTINUA.†

By J. H. Roberts.‡

The notion *upper semi-continuous collection of bounded continua,* introduced by R. L. Moore,§ has been very fruitful in the realm of analysis situs. In particular, Moore has shown that if G is an upper semi-continuous collection of bounded continua¶ filling a plane S, then if no element of G separates S the collection G is itself homeomorphic with S.‖ If the restriction that no element of G separates S is removed then G is †† homeomorphic with some cactoid.††

I have attempted to extend the idea *upper semi-continuous collection* to apply to collections whose elements were not all bounded. In my paper ‡‡ "Concerning Collections of Continua not all Bounded," I suggest a definition. I there show that if G is an upper semi-continuous *and metric* collection of continua filling a plane S and no element of G separates S, then G is homeomorphic with a *subset* of some cactoid.§§ In the present paper it is shown that without the restriction that no element of G separates S, G is homeomorphic with a subset of some cactoid.

Definition. If $h_1, h_2, h_3, \cdots$ denotes any sequence of elements of a collection G of continua, and for each i the element h_i contains points P_i and Q_i then, granting that the sequence $P_1, P_2, P_3, \cdots$ has a sequential limit point P in an element g_P of G, the collection G is said to be *upper semi-continuous* provided every limit point of $\sum_{i=1}^{\infty} Q_i$ lies in g_P.

---

† Presented to the Society, under a different title, June 21, 1929.

‡ National Research Fellow.

§ "Concerning Upper Semi-Continuous Collections of Continua," *Transactions of the American Mathematical Society*, Vol. 27 (1925), pp. 416-428.

¶ Throughout this paper a *point* is considered as a special type of *continuum.*

‖ With respect to a suitable definition of limit element. See Moore, *loc. cit.*

†† Moore, "Concerning Upper Semi-continuous Collections," *Monatshefte für Mathematik und Physik*, Vol. 36 (1929), pp. 81-88. A *cactoid* is a bounded continuous curve M lying in space of three dimensions and such that (a) every non-degenerate maximal cyclic subset of M is a simple closed surface and (b) no point of M lies in a bounded complementary domain of any subcontinuum of M.

‡‡ *American Journal of Mathematics*, Vol. 52 (July, 1930). Hereafter this paper will be referred to as C. C.

§§ In fact, except for the case where exactly one element of G is unbounded, G is homeomorphic with a subset of the plane.

Definition. If G is an upper semi-continuous collection of continua, then an element g of G is said to be a *limit element* of a set K of elements of G if g contains a point P which is a limit point of the point set obtained by adding together all elements of K except g.

The preceeding definitions apply to collections all of whose elements lie in a space S for which the notion *limit point* has been defined. Note that the elements of an upper semi-continuous collection of continua are mutually exclusive.

THEOREM 1. *If G is an upper semi-continuous and metric collection of continua filling a continuous curve* † *M, then G is itself a continuous curve.*

Proof. It is easily seen that G is separable, connected, and locally connected. I shall show that it is also locally compact. Let g be any element of G. There exists a sequence of points $P_1, P_2, P_3, \cdots$ lying in g such that every point of g is a limit point of $\sum_{i=1}^{\infty} P_i$. Let d and δ denote distance functions for the spaces M and G, respectively. I shall define, for each i, a number d_i. If the M-domain $S(P_i, 1/2)$ ‡ is compact, then let d_i be 1/2. Otherwise, let d_i be 1/2 the upper limit of d where $S(P_i, d)$ is compact. Now let K_n be the point set $\sum_{i=1}^{n} S(P_i, d_i)$, and let R_n be the set of elements of G which contain at least one point of K_n. For every n the set R_n is compact and contains g. If for every n the element g is a limit element of $G - R_n$, then there is a sequence of elements $g_1, g_2, g_3, \cdots$ such that $\delta(g, g_n) < 1/n$, and g_n does not belong to R_n. Then as limit of $\delta(g, g_n) = 0$ there is, for each n, a point Q_n in g_n such that $\sum_{i=1}^{\infty} Q_i$ has a limit point Q in g. Let D be a compact M-domain containing the point Q. Obviously as M is metric there is an integer n such that $S(P_n, d_n)$, and therefore K_n, contains Q. But as K_n is a domain containing a limit point Q it must contain, for some m greater than n, the point Q_m. Then g_m belongs to K_m, and we have reached a contradiction, which shows that for some n the element g is not a limit element of the set $G - R_n$. Thus the compact set R_n contains a domain containing g.

THEOREM 2. *If G is an upper semi-continuous and metric collection of*

† In the present paper a space M is said to be a continuous curve if it is metric, connected, locally connected, separable, and locally compact.

‡ That is, the set of all points of M whose distance from P_i is less than 1/2.

continua filling a plane S, then each cyclic element † *of the continuous curve G is homeomorphic with a subset of a sphere.*

Proof. Let M be any non-degenerate cyclic element of G. If g is an element of M it follows, since $M - g$ is connected, that just one complementary domain E_g of g in the plane S contains points belonging to elements of $M - g$. Let g^* be $S - E_g$. If g_1 and g_2 are distinct elements of M then g_1^* and g_2^* are distinct. Let M^* be the collection of continua containing g^* for every element g of M. Now M^* fills the plane S, for if P is a point of S and g_P—the element of G containing P—does not belong to M, then since M is a cyclic element of G, some element g of M separates g_P and $M - g$ in S. Thus g_P is a subset of g^*. It follows that M^* is an upper semi-continuous and metric collection of continua, homeomorphic with M, and such that no element of M^* separates S, and every point of S belongs to some element of M^*. Then by theorems 3 and 4 of C. C. it follows that M^*, and therefore M, is homeomorphic with a subset of a sphere.

THEOREM 3. *If every cyclic element of a continuous curve G is homeomorphic with some subset of a sphere, then G itself is homeomorphic with a subset of some cactoid.*

Proof. Let E_1 and E_2 be distinct cyclic elements of G, and let X be a simple cyclic chain ‡ of G between E_1 and E_2. Let A_1A_2 denote an arc such that A_i lies in E_i, but does not lie in any other cyclic element of G except when E_i is degenerate $(i = 1, 2)$. Let K be the set of points of G which separate A_1 and A_2. Then by a theorem of Whyburn's,§ for every maximal segment s of $A_1A_2 - (K + A_1 + A_2)$ there is a unique cyclic element of G containing s. Let $s_1, s_2, s_3, \cdots$ denote the maximal segments of $A_1A_2 - (K + A_1 + A_2)$. Let π be a continuous one to one correspondence throwing A_1A_2 onto the interval $0 \leqq x \leqq 1$, $y = 0$, $z = 0$. For each i let F_i be a sphere with $\pi(\bar{s}_i)$ as diameter. Then the set M $[M = \sum_{i=1}^{\infty} F_i + (K + A_1 + A_2)]$

† By a *cyclic element* of a continuous curve M is meant (a) a maximal cyclic curve of M, (b) a cut point of M, or (c) an end point of M. See G. T. Whyburn, "Concerning the Structure of a Continuous Curve," *American Journal of Mathematics*, Vol. 50 (1928), pp. 167-194.

‡ See Whyburn, *loc. cit.* A point set X is said to be a *simple cyclic chain* of G between E_1 and E_2 provided that X is connected and contains E_1 and E_2 and is the sum of the elements of some collection of the cyclic elements of G, and furthermore no proper connected subset of X containing E_1 and E_2 is the sum of the elements of such a collection.

§ "Some Properties of Continuous Curves," *Bulletin of the American Mathematical Society*, Vol. 33 (1927), pp. 305-308, theorem II.

is a cactoid. Moreover, as any two distinct points of a sphere can be thrown into any two distinct points of the sphere by a continuous transformation of the sphere into itself, it follows that the cyclic chain X is homeomorphic with a subset of the cactoid M.

Now it is easily seen that there exists a sequence $Q_1, Q_2, Q_3, \cdots$ of distinct cyclic elements of G such that every point of G is a limit point of $\sum_{i=1}^{\infty} Q_i$. Now for every m and n $(m \neq n)$ G contains a unique simple cyclic chain between Q_m and Q_n. If i, n_i, and m_i are positive integers, let C_i denote the simple cyclic chain of G between Q_{n_i} and Q_{m_i}. Let $n_1 = 1$ and let $m_1 = 2$. For i greater than 1 the integers n_i and m_i will be defined by induction as follows: Having defined integers $n_1, n_2, \cdots, n_i$ and $m_1, m_2, \cdots, m_i$ $(i \geqq 1)$ let n_{i+1} be the smallest integer k such that Q_k does not belong to $\sum_{j=1}^{i} C_j$. Let m_{i+1} be the integer k such that Q_k is the first element of the chain from $Q_{n_{i+1}}$ to Q_{n_1} which belongs to the set $\sum_{j=1}^{i} C_j$. Clearly every point of G not an end point belongs, for some i, to the chain C_i. It has been shown that every simple cyclic chain of G is homeomorphic with a subset of some cactoid.

It can be shown that there exist cactoids $D_1, D_2, D_3, \cdots$, and correspondences $\pi_1, \pi_2, \pi_3, \cdots$ such that (1) π_i is a continuous one to one correspondence between $\sum_{j=1}^{i} C_j$ and a subset of D_i, (2) if P is a point of C_i then $\pi_i(P) = \pi_{i+1}(P) = \pi_{i+2}(P) = \cdots$, (3) the cactoid D_i is a subset of D_{i+1}, (4) for every i and k $(i < k)$ the components of $D_k - D_i$ can be ordered $H_1, H_2, \cdots, H_n$ so that if P_j $(j \leqq n)$ denotes the limit point of H_j which belongs to D_i, and K_j denotes the sum of the sets $\bar{H}_1, \bar{H}_2, \cdots, \bar{H}_{j-1}$ which do not contain P_j, then the diameter of H_j is less than the smaller of the numbers $1/4^i$ and $d(P_j, K_j)/4^i$, and (5) if $i < j < k$, H is a component of $D_k - D_j$, and P is the limit point of H which belongs to D_j, then if P is not in D_i the diameter of H is less than $d(P, D_i)/4^i$.

Let D denote $\sum_{i=1}^{\infty} D_i$. From property 4 above it follows that the continuum D is a continuous curve. If T is a point which belongs to D_i and is a limit point of $D_{i+1} - D_i$, then for every k $(k > i)$ let g_{Tk} denote any component of $D_k - D_i$ with T as limit point. Then (see properties 4 and 5) the set $g_{Tk} + g_{T(k+1)} + g_{T(k+2)} + \cdots$ has no limit point in D_i except T. Moreover, if T_1 and T_2 are distinct points which belong to D_i and are limit points of $D_k - D_i$ $(k > i)$ then no point is a limit point of both the sets $g_{T_1k} + g_{T_1(k+1)} + g_{T_1(k+2)} + \cdots$ and $g_{T_2k} + g_{T_2(k+1)} + g_{T_2(k+2)} + \cdots$. It fol-

lows that every simple closed curve belonging to $\bar{D}$ belongs, for some i, to some topological sphere of the cactoid D_i. Obviously no point lies within any sphere of $\bar{D}$. Thus $\bar{D}$ is a cactoid.

Let P be a point of G not belonging, for any i, to C_i, and let P_1, P_2, $P_3, \cdots$ denote a sequence of points converging to P such that, for each i, P_i belongs to $\sum_{j=1}^{i} C_j$. Obviously for every i the points of a subsequence V of $P_1, P_2, P_3, \cdots$, where V contains all but a finite number of points of the original sequence, belong to a single component of $G - \sum_{j=1}^{i} C_j$. Then the sequence $\pi_1(P_1)$, $\pi_2(P_2)$, $\pi_3(P_3), \cdots$ converges to a point P'. Moreover P' is uniquely determined by the point P. If P is a point of C_i then let P' be $\pi_i(P)$. Thus for every point P of G there has been defined a point P' belonging to $\bar{D}$. The correspondence π which is such that $\pi(P) = P'$ is a continuous one to one correspondence between G and a subset of the cactoid $\bar{D}$. Thus theorem 3 is established.

COROLLARY. *If G is any upper semi-continuous and metric collection of continua filling a plane, then G is homeomorphic with a subset of some cactoid.*

THEOREM 4. *If G is an upper semi-continuous and metric collection of continua filling a plane S, and is such that* (1) *no element of G which is a bounded continuum in S separates S, and* (2) *every maximal domain D of bounded elements of G has at least two elements of G on its boundary, then G is homeomorphic with a subset of the plane.*

Proof. From theorem 1 it follows that the hyperspace G is a continuous curve each of whose cyclic elements is homeomorphic with a subset of a sphere. But in view of the hypotheses of the present theorem, and theorem 3 of C. C., it is clear that no cyclic element of the hyperspace G is a sphere. Thus each cyclic element of G is homeomorphic with a subset of a plane. Now no element g of G which is a bounded continuum in S separates S. Then g does not separate G. Hence if M is a cyclic element of G, those elements of M which are limit elements of $G - M$ must be unbounded continua in S. But a correspondence π throwing M into a set as described in theorem 4 of C. C. throws every unbounded element of G into a point which is arcwise accessible from the complement of $\pi(M)$. Thus theorem 4 can be proved by a modification of the proof of theorem 3.

THE UNIVERSITY OF PENNSYLVANIA.

THE CYCLIC AND HIGHER CONNECTIVITY OF LOCALLY CONNECTED SPACES.

By G. T. Whyburn.

1. *Introduction.* In this paper we shall consider connected and locally connected, separable metric spaces which we shall denote by the letter M. By a *region* in such a space is meant any connected open set of points. The point set $M - \bar{R}$ will be called the exterior of the region R. A region will be said to *join* two point sets A and B provided that $A \cdot R \neq 0 \neq B \cdot R$. Two regions R and S are said to be *strongly separated* provided that $\bar{R} \cdot \bar{S} = 0$. The point p is a cut point of a space M provided $M - p$ is not connected. A point x is called an end point of M provided there exists a monotonic decreasing sequence $[U_i]$ of neighborhoods of x, such that the boundary of U_i is a single point and such that $x = \prod_1^\infty \bar{U}_i$. A subset N of M is called a *nodule* of M if N is a connected subset of M which contains more than one point, has no cut point and is saturated in M with respect to these properties. The cut points, end-points, and nodules of a space M are called the *nodular elements* of M.*

If R is a region, $F(R)$ will denote the boundary of R, i. e., the set of points $\bar{R} - R$. For each point p and each positive number r, $S_r(p)$ and $V_r(p)$ will denote respectively the set of all points whose distances from p are equal to and less than r; $U_r(p)$ will denote the component of $V_r(p)$ which contains p.

This paper is devoted to the study of what might be termed, broadly speaking, the connectivity properties of spaces M and their relation to similar known properties of continuous curves, i. e., spaces M which are locally compact. The results found effect, in many cases, real generalizations of the known connectivity properties of continuous curves to the more general spaces M. Before proceeding with the theorems, however, we give in the next section an example which illustrates the complexity which may arise even in a greatly restricted space M.

* For a discussion of the structure of spaces M with respect to their nodular elements, see the author's paper in the *Transactions of the American Mathematical Society*, Vol. 32 (1930), pp. 926-943.

2. *Example.* There exists a space M which has no cut point, is both an absolute G_δ* and an absolute F_σ*, and yet is not cyclicly connected.*

Let a and b respectively denote the points $(1, 0)$ and $(-1, 0)$, and let ab denote the interval $(-1, 1)$. For each positive integer n let C_n denote the semi-circle $y = [(1 - 1/n)^2 - x^2]^{1/2}$. Let $\mathfrak{M} = ab + \sum_1^\infty C_n$. Then clearly $\mathfrak{M}$ has all the desired properties and yet a and b lie together on no simple closed curve in $\mathfrak{M}$.

Thus it is seen that although the property of arcwise connectivity of continuous curves extends † also to spaces M which are absolute G_δ's, the cyclic connectivity property of continuous curves without cut points does not extend to absolute G_δ spaces M without cut points, even though M is at the same time an absolute F_σ.

It is interesting to note that even in this example $\mathfrak{M}$ there does exist two mutually exclusive regions R and S such that $\bar{R} \cdot \bar{S} = a + b$.

3. *The quasi-divisibility of spaces M.*

(3.1). Theorem. *If p is any non-cut point of M, there exists a monotone decreasing sequence $R_1, R_2, R_3, \cdots$ of regions such that $p = \prod_1^\infty \bar{R}_i$, $\lim_{i\to\infty} \delta[p + F(R_i)] = 0$ and for each i the exterior G_i of R_i is connected and $F(G_i) = F(R_i)$.*

Proof. Let q be any point of $M - p$. For each $r < \rho(p, q)$, let $G_r(q)$ be the component of $M - \overline{U_r(p)}$ containing q, let $X_r = F[G_r(q)]$, and let $R_r(p)$ be the component of $M - X_r$ containing p. Then clearly we have

$$\text{(i)} \qquad X_r = F[G_r(q)] = F[R_r(p)] \subset F[U_r(p)] \subset S_r(p).$$

Now if $r_1 < r_2 < \rho(p, q)$, we have

$$\text{(ii)} \qquad G_{r_2}(q) + X_{r_2} = \overline{G_{r_2}(q)} \subset G_{r_1}(q),$$

and therefore

$$\text{(iii)} \qquad R_{r_1}(p) \subset R_{r_2}(p),$$

for $R_{r_1}(p) \cdot X_{r_2} = 0$ by (ii).

---

* A G_δ set is one which is the product of some family of open sets. An F_σ set is one which is the sum of a countable number of closed sets. A G_δ or F_σ set is called *absolute* provided it is a G_δ or F_σ in every metric space in which it is topologically contained. A space M is said to be *cyclicly connected* if every two points of M lie together on a simple closed curve of M.

† See R. L. Moore, *Bulletin of the American Mathematical Society*, abstract, Vol. 33 (1927), p. 141; also K. Menger, *Monatshefte für Mathematik und Physik*, Vol. 36 (1930), p. 210.

Now for all save a countable number of r's, we have $R_r(p) + X_r + G_r(q) = M$. For let E be the set of all numbers $e < \rho(p, q)$ such that a component D_e of $M - X_e$ other than $R_e(p)$ and $G_e(q)$ exists. Then for $e_1 < e_2$ it follows by (ii) and (iii) since $F[D_{e_2}] \subset X_{e_2}$ that $D_{e_2} \subset G_{e_1}(q)$ and therefore that $D_{e_1} \cdot D_{e_2} = 0$. Hence E must be countable, for the sets D_e are open. Thus there exists a sequence of numbers $r_1, r_2, r_3, \cdots$, converging monotonically to zero such that for each i, $M = R_{r_i}(p) + X_{r_i} + G_{r_i}(q)$. Now if x is any point of $M - p$, there exists a connected set I containing $q + x$ such that $\bar{I} \cdot p = 0$, for p is not a cut point of M. Hence there exists an integer i such that $q + x \subset I \subset G_{r_i}$. Therefore $p = \prod_1^\infty \bar{R}_{r_i}(p)$. Thus if for each i, we set $R_i = R_{r_i}(p)$, $G_i = G_{r_i}(q)$, the sets $[R_i]$ and $[G_i]$ have all the necessary properties for the theorem. $\lim_{i \to \infty} \delta[p + F(R_i)] = 0$, because

$$p + F(R_i) = p + X_{r_i} \subset \overline{V_{r_i}(p)} \text{ and } \delta[V_{r_i}(p)] \leq 2r_i.$$

Corollary (3.1a). *If H is the sum of a finite number of connected sets and $\bar{H} \cdot p = 0$, then there exists an integer i such that $\bar{H} \subset G_i$ and hence $\bar{H} \cdot \bar{R}_i = 0$.*

Corollary (3.1b). *The regions $R_1 = R_{r_1}(p)$, $R_2 = R_{r_2}(p), \cdots$, may be chosen so that for every n, $| r_n - 1/n | < 1/n^2$; or more generally, if $a_1, a_2, \cdots$ is any sequence of numbers $< \rho(p, q)$ which converges monotonically to zero, then, so that for each n, $| r_n - a_n | < 1/n^2$.*

Corollary (3.1c). *If M is locally compact, then each non-cut point of M is contained in arbitrarily small regions whose exteriors (also whose complements) are connected.**

Note. Corollary (3.1a) yields a very simple proof for the theorem † that the set G of all non-cut points of a space M is a G_δ set (or that the set F of all cut points is an F_σ set). For let $\sum_1^\infty p_i$ be a countable set of points dense in M. For each $n > 0$, let G_n be the set of all points x of M such that x lies in some region R whose exterior is connected and contains every point

* For theorems closely related to this corollary, see H. M. Gehman, *Proceedings of the National Academy of Sciences*, Vol. 14 (1928), pp. 431-433, and W. L. Ayres, *Monatshefte für Mathematik und Physik*, Vol. 36 (1929), pp. 139-140. Indeed, the two theorems in this section of the present paper may be regarded as generalizations of the theorems of Gehman–Ayres to be found in these references.

† See G. T. Whyburn, *loc. cit.*, result (1.3). For the case of compact spaces M, see Zarankiewicz, *Fundamenta Mathematicae*, Vol. 9, p. 163. For a simple proof of Zarankiewicz's theorem, see W. L. Ayres, *Fundamenta Mathematicae*, Vol. 16.

except possibly one of $\sum_1^n p_i$. Then clearly G_n is open, and it follows at once from Corollary (3.1a) that $G = \prod_1^\infty G_n$.

(3.2) THEOREM. *If p is any point of a space M, there exists a monotone decreasing sequence $R_1, R_2, \cdots$ of regions such that $p = \prod_1^\infty \bar{R}_i$, $\lim_{i\to\infty} \delta[p + F(R_i)] = 0$ and for each i the exterior G_i of R_i is the sum of a finite number, n_i, of connected sets and $F(G_i) = F(R_i)$. Furthermore n_i does not exceed either i or the number of the components of $M - p$, and no two of these n_i sets lie together in the same component of $M - p$.*

Proof. Let $M_1, M_2, M_3, \cdots$ be the components of $M - p$. For each n the set $M_n + p$ is connected and locally connected and has p for a non-cut point. By virtue of (3.1) and its proof it follows that there exists a sequence of number $d_1, d_2, \cdots$ converging monotonically to zero and for each n there exists a sequence of regions (relative to $M_n + p$) $R_1{}^n, R_2{}^n, \cdots$ such that $p = \prod_{i=1}^\infty \bar{R}_i{}^n$ and for each i the exterior, $G_i{}^n$, in $M_n + p$ of $R_i{}^n$ is connected and the boundary $X^n{}_{d_i}$ of $R_i{}^n$ and $G_i{}^n$ is a subset of $S_{d_i}(p)$. Now set $R_1 = M - \bar{G}_1{}^1$, $R_2 = M - (\bar{G}_2{}^1 + \bar{G}_2{}^2), \cdots$, $R_n = M - \sum_{i=1}^n \bar{G}_n{}^i, \cdots$, it being understood that if there are only a finite number, k, of components of $M - p$, then for all n's $> k$, and all i's, $M_n = R_i{}^n = G_i{}^n = X_{d_i}{}^n = 0$. Then for each n, $F(R_n) = \sum_{i=1}^n X_{d_n}{}^i \subset S_{d_n}(p)$ and $M - \bar{R}_n = G_n = \sum_{i=1}^n G_n{}^i$; and it follows at once that the sets $R_1, R_2, \cdots$ satisfy all the required conditions in the theorem.

COROLLARY (3.2). *If p is any point of a space M and H is a closed subset of $M - p$ which is the sum of a finite number of connected sets, then there exists a region R containing p whose exterior contains H and is the sum of a finite number of connected sets.*

4. *Locally divisible spaces M.*

Definition. A space M will be said to be *locally divisible in the point p* provided that p is contained in arbitrarily small regions whose exteriors are connected or are the sum of a finite number of connected sets according as p is a non-cut point or a cut point of M; a space which is locally divisible in each of its points will be said to be *locally divisible*.†

† It should be noted that this notion is a localization of a stronger property than that of *divisibility* as introduced by W. A. Wilson. See Wilson, *Bulletin of the American Mathematical Society*, Vol. 36 (1930), p. 85.

Examples are easily constructed of spaces M which are not locally divisible—indeed which are not locally divisible in any one of their points.

(4.1) THEOREM. *In order that M be locally divisible in p it is necessary and sufficient that p be contained in arbitrarily small regions whose complements are connected or are the sum of a finite number of connected sets.*

(4.2) THEOREM. *Any space M may be transformed by a biunivalued and continuous transformation into a connected, locally connected, separable metric space M^* which is locally divisible.*

Proof. Let G denote the collection of all regions in M whose exteriors and whose complements are the sum of a finite number of connected sets. Let M^* denote the space whose points are exactly the points of M but in which limit point is defined by means of the system of neighborhoods G, i. e., the point p^* of M^* is a limit point of the point set N^* in M^* if and only if every region of the collection G which contains p^* contains at least one point of N^* distinct from p^*. Now since every set E^* in M^* which is open in M^* is also an open set E in M, it follows that M^* may be regarded as the image of M under a biunivalued and continuous transformation T. It remains to show that M^* is connected, locally connected, separable, metric, and locally divisible. Now M^* is connected, locally connected and separable because it is the image under T of M, and M has these properties. To show that M^* is metric we need only prove that it is regular and perfectly separable. It is regular, because if p^* is any point of M^* and U^* is any neighborhood of p^*, there exists a region E^* of G containing p^* and lying in U^*. Since $M - E$ is the sum of a finite number of connected sets, there exists, by corollary (3.2), a region V of G containing p whose exterior contains $M - E$. Therefore, $\bar{V} \subset E$; whence $\bar{V}^* \subset E^*$, which proves M^* regular. To show that M^* is perfectly separable, let $p_1, p_2, p_3, \cdots p_n \cdots$ be a countable sequence of points dense in M. Arrange the positive rational number into a sequence $r_1, r_2, r_3, \cdots r_i \cdots$. For each n and each i, arrange all possible point sets Q such that Q is the sum of a finite number of the components of $M - \overline{V_{r_i}(p_n)}$ into a sequence $Q_1(i, n), Q_2(i, n), \cdots, Q_j(i, n), \cdots$. For each j, i, and n and for each point x of $U_{r_i}(p_n)$, there exists, by corollary (3.2), a region G_x of the collection G containing x but containing no point of $Q_j(i, n)$. By the Lindelöf theorem there exists a countable collection $[G_k(j, i, n)]_{k=1}^{\infty}$ of the regions G_x whose sum covers $U_{r_i}(p_n)$. Let E denote the collection of regions $[G_k(j, i, n)]$, for all n's, i's, j's and k's. Then E is countable and is equivalent to the system G. For let p be any point and let R_g be any region of G

containing p. There exists an i such that $V_{4r_i}(p) \subset R_g$. There exists an n such that $p_n \subset U_{r_i}(p)$ and hence such that

$$U_{r_i}(p) \subset U_{2r_i}(p_n) \subset V_{2r_i}(p_n) \subset V_{4r_i}(p) \subset R_g.$$

Then since $M - R_g$ is the sum of a finite number of connected sets, there exists a j such that $Q_j(i', n) \supset M - R_g$, where i' is an integer such that $r_{i'} = 2r_i$. There exists a k so that $G_k(j, i', n)$ contains p but contains no point of $Q_j(i', n)$ and hence no point of $M - R_g$. Thus we have

$$p \subset G_k(j, i', n) \subset R_g,$$

which proves E equivalent to G, inasmuch as E is a subcollection of G. Therefore M^* is a perfectly separable, regular, Hausdorff space and is then metrisable by the theorem of Alexandroff–Urysohn–Tychonoff.† We suppose a distance function defined in M^*. Clearly M is locally divisible, because for each ϵ the set $V^*_\epsilon(p^*) = V^*$ is open in M^* for each point p^* in M^*; and thus there exists a region R_g of the collection G which contains p and lies in V. Hence $p^* \subset R^*_g \subset V^*_\epsilon(p^*)$. Whence $\delta(R^*_g) < 2\epsilon$, and $M^* - \bar{R}^*_g$ is the sum of a finite number of connected sets.

5. *Local end points.*

Definition. A point p which is an end point of at least one region in M is called a *local end point* of M. It is obvious that if M is compact, its local end points are identical with its end points. For general M-spaces, however, this is not true. For example, the points a and b of the space $\mathfrak{M}$ described in § 3 are local end points but not end points of $\mathfrak{M}$.

(5.1) LEMMA. *If in a space M having no cut point, A, B_1 and B_2 are mutually exclusive closed sets such that A is non-degenerate, and there exist regions G_1 and G_2 joining A and B_1, and A and B_2 respectively such that $G_1 \cdot B_2 = G_2 \cdot B_1 = 0$, then there exist two strongly separated regions R_α and R_β joining A and B_1, and A and B_2 respectively, such that $R_\alpha \cdot B_2 = R_\beta \cdot B_1 = 0$.*

Proof. Let S_1 denote the set of all points x of M such that strongly separated regions R_a and R_x exist joining A and B_1, and x and B_2 respectively, such that $R_a \cdot B_2 = R_x \cdot B_1 = 0$; similarly let S_2 be the set of all points x of M such that strongly separated regions R_a and R_x exist joining A and B_2, and x and B_1 respectively, and such that $R_a \cdot B_1 = R_x \cdot B_2 = 0$. It follows at once from our hypothesis that the sets S_1 and S_2 are non-vacuous and contain B_1 and B_2 respectively, and that their sum S is an open set. We

† See articles by these authors in *Mathematische Annalen*, Vols. 92-95.

shall show that $S = M$. If this is not so, then there exists at least one point y of $M - S$ which is a limit point of S. Since y is not a cut point of M, since A contains more than one point, and y does not belong to $B_1 + B_2$, it follows that there exist strongly separated regions R^* and G such that R^* joins A and $B_1 + B_2$ and G contains y but $\bar{G} \cdot (B_1 + B_2) = 0$. The region G contains at least one point x of S, and thus there exist strongly separated regions R_a and R_x such that either R_a joins A and B_1 but contains no point of B_2 and R_x joins x and B_2 and contains no point of B_1, or R_a joins A and B_2, R_x joins x and B_1, and $R_a \cdot B_1 = R_x \cdot B_2 = 0$. The two cases are entirely alike, so we shall suppose the former. Let R_1 and R_2, respectively, denote components of $R_a - R_a \cdot \bar{G}$ and $R_x - R_x \cdot \bar{G}$ containing points of B_1 and B_2 respectively. Then since y does not belong to S, it follows at once that $\bar{R}_1 \cdot \bar{G} \neq 0 \neq \bar{R}_2 \cdot \bar{G}$. Let $K = \bar{R}_1 + \bar{R}_2 + B_1 + B_2$, and let R denote a component of $R^* - R^* \cdot K$ which contains at least one point of A. Then K contains at least one limit point p of R. Either $p \subset \bar{R}_1 + B_1$ or $p \subset \bar{R}_2 + B_2$. Here again the two cases are alike, so we shall treat only the former. Let q denote a point of $\bar{R}_2 \cdot \bar{G}$, and let U_p and U_q be strongly separated regions, containing p and q respectively, such that $\bar{U}_p \cdot (\bar{R}_2 + \bar{B}_2 + \bar{G}) = 0$ and $\bar{U}_q \cdot (\bar{R}_1 + B_1 + \bar{R}) = 0$. Clearly $R + U_p + R_1$ contains a region W joining A and B_1, and $R_2 + U_q + G$ is a region, say Q, joining y and B_2 and we have $W \cdot Q = 0$, $W \cdot B_2 = Q \cdot B_1 = 0$. Now obviously W contains a region V joining A and B_1 and such that $\bar{V} \subset W$. But then Q and V are strongly separated and join A and B_1, and y and B_2, respectively, and $V \cdot B_2 = Q \cdot B_1 = 0$, contrary to the fact that y does not belong to S. Thus the supposition that $S \neq M$ leads to a contradiction. Accordingly, S contains a point s of A, and thus there exist two strongly separated regions $R_a = R_\alpha$ and $R_x = R_\beta$ joining A and B_1, and A and B_2, respectively, such that $R_\alpha \cdot B_2 = R_\beta \cdot B_1 = 0$.

Corollary (5.1a). *If the space M as in the Lemma is imbedded in a space M_0, also an M-space, and if B_1^0 and B_2^0 are subsets of M_0 such that $\bar{B}_1^0 \cdot M = B_1$ and $\bar{B}_2^0 \cdot M = B_2$ and ϵ is any positive number, then there exist strongly separated regions R_1 and R_2 in M_0 joining A and B_1, and A and B_2, respectively, such that $R_1 \cdot B_2^0 = R_2 \cdot B_1^0 = 0$ and every point of $R_1 + R_2$ is at a distance $< \epsilon$ from some point of M.*

Corollary (5.1b). *If H and K are non-degenerate mutually exclusive subsets of a space M which has no cut point, then there exist two strongly separated regions in M joining H and K.*

(5.2). Theorem. *If the point p is not a local end point of a space M,*

then there exist two regions R_1 and R_2 such that $R_1 \cdot R_2 = p$ and the sets $R_1 + p$ and $R_2 + p$ are connected and locally connected.

Proof. We may suppose p is neither a cut point nor a local separating point of M, for in these cases the theorem is obvious. Hence there exists a nodule N of M containing p. Let a and b be any two points of $N - p$ and let U_1 denote the component containing p of $N \cdot V_{\epsilon_1}(p)$, where $\epsilon_1 < 1$ and also $\epsilon_1 < \rho(p, a + b)$. Likewise, we may suppose p is neither a cut point nor an end point of U_1, for it is seen at once that if a point p is a local end point of some nodule of a given space M, then p is a local end point of the whole space. Hence there exists a nodule N_1 of U_1 which contains p. Since N has no cut point, then by virtue of the preceding lemma and its corollaries there exist two strongly separated regions S^*_a and S^*_b in M joining a and N_1, and b and N_1, respectively. Let S_a and S_b, respectively, denote the components of $S^*_a - \bar{N}_1 \cdot S^*_a$ and $S^*_b - \bar{N}_1 \cdot S^*_b$ which contain a and b respectively. Now in case $\bar{S}_a \cdot N_1$ and $\bar{S}_b \cdot N_1$ or either is non-vacuous, then let A_1 and B_1 respectively denote single points of these sets. All cases which may arise here, however, are only simpler than that in which these two sets are vacuous; hence we shall suppose that $\bar{S}_a \cdot N_1 = \bar{S}_b \cdot N_1 = 0$. Then let a_1 and b_1 respectively denote points of $\bar{S}_a \cdot \bar{N}_1$ and $\bar{S}_b \cdot \bar{N}_1$; and let G_{a_1} and G_{b_1} denote strongly separated regions in M of diameter $< 1/2$ containing a_1 and b_1 respectively and such that $\bar{G}_{a_1} \cdot (\bar{S}_b + p) = \bar{G}_{b_1} \cdot (\bar{S}_a + p) = 0$. Let U_2 be a region containing p and of diameter $< 1/2$ such that $\bar{U}_2 \cdot \overline{(S_a + S_b + G_{a_1} + G_{b_1})} = 0$, and let N_2 be the (necessarily non-degenerate) nodule of U_2 which contains p. Now let N^*_1 be the nodule of the set $N_1 + G_{a_1} + G_{b_1}$ which contains N_1 and let $A_1 = \bar{S}_a \cdot N^*_1$ and $B_1 = \bar{S}_b \cdot N^*_1$. Let $A_1{}^0 = S_a$, $B_1{}^0 = S_b$. Then A_1 and B_1 are mutually exclusive closed subsets of N^*_1 and they contain a_1 and b_1 respectively but contain no points of $\bar{N}_2$. Also A_1 does not separate N_2 and B_1 in N^*_1 nor does B_1 separate N_2 and A_1 in N^*_1; for there exists a point x of N_1 in G_{b_1} which can be joined to N_2 by a region R in N_1 having no limit points in A_1, and hence $R + G_{b_1} \cdot N^*_1$ is a connected set in N^*_1 joining N_2 and B_1 and having no limit points in A_1, and similarly for the sets N_2 and A_1. Hence by the lemma and corollary (5.2b), there exist in M two strongly separated regions $S^*_{a_1}$ and $S^*_{b_1}$ joining N_2 and A_1, and N_2 and B_1, respectively, and such that $B_1{}^0 \cdot \bar{S}^*_{a_1} = A_1{}^0 \cdot \bar{S}^*_{b_1} = 0$, and such that every point of these two regions is at a distance $< 1/2$ from some point of N^*_1, and hence so that $\bar{S}^*_{a_1} \cdot \bar{S}_b = \bar{S}^*_{b_1} \cdot \bar{S}_a = 0$. Let S_{a_1} and S_{b_1}, respectively, denote components of $S^*_{a_1} - \bar{N}_2 \cdot S_{a_1}$ and $S^*_{b_1} - \bar{N}_2 \cdot S^*_{b_1}$ containing points of A_1 and B_1 respectively. Again we shall suppose, without real loss of generality, that $\bar{S}_{a_1} \cdot N_2 = \bar{S}_{b_1} \cdot N_2 = 0$. Let a_2 and b_2 be points of $\bar{S}_{a_1} \cdot \bar{N}_2$ and $\bar{S}_{b_1} \cdot \bar{N}_2$ respec-

tively, and let G_{a_2} and G_{b_2} be strongly separated regions in M containing a_2 and b_2 respectively, each of diameter $< 1/3$, and such that

$$\bar{G}_{a_2} \cdot (\bar{S}_b + \bar{S}_{b_1} + p) = \bar{G}_{b_2} \cdot (\bar{S}_a + \bar{S}_{a_1} + p) = 0.$$

Let U_3 be a region of diameter $< 1/3$ containing p and such that

$$\bar{U}_3 \cdot \overline{(S_a + S_{a_1} + S_b + S_{b_1} + G_{a_2} + G_{b_2})} = 0,$$

and let N_3 be the nodule of U_3 which contains p. Then just as above it follows that there exist in M two strongly separated regions $S^*_{a_2}$ and $S^*_{b_2}$ joining S_{a_1} and N_3, and S_{b_1} and N_3, respectively, and such that

$$\bar{S}^*_{a_2} \cdot \overline{(S_b + S_{b_1})} = \bar{S}^*_{b_2} \cdot \overline{(S_a + S_{a_1})} = 0,$$

and every point of $S^*_{a_2} + S^*_{b_2}$ is at a distance $< 1/3$ from some point of N^*_2, where N^*_2 is the nodule of the set $N_1 + G_{a_2} + G_{b_2}$ containing N_2. Let S_{a_2} and S_{b_2} be components, $\cdots$, and so on. Continue this process indefinitely. Let $R_1 = S_a + \sum_1^\infty S_{a_n}$ and $R_2 = S_b + \sum_1^\infty S_{b_n}$. Then R_1 and R_2 have the desired properties for the theorem. For, every point of $\sum_n^\infty (S_{a_i} + S_{b_i})$ is at a distance $\leq 3/n$ from p and hence it follows that $\bar{R}_1 \cdot \bar{R}_2 = p$ and that $R_1 + p$ and $R_2 + p$ are locally connected.

COROLLARY (5.2). *In order that the point p of a space M should be a local end point of M, any one of the following conditions is necessary and sufficient:*

1) *that there should not exist two regions R_1 and R_2 such that $R_1 \cdot \bar{R}_2 = p$ and $R_1 + p$ and $R_2 + p$ are locally connected.*

2) *that there should not exist two mutually exclusive simply infinite chains* † *of regions each converging to p.*

3) *that p should not be a cut point of any connected and locally connected subset of M.*

6. *A Characterization of end points.*

Definition. If a and p are points of a space M, the subset C of M will be called a *simply infinite chain* of regions from a converging to p provided that C is the sum of a sequence of regions $R_1, R_2, \cdots$, where $R_1 \supset a$, $C \cdot p = 0$, for $i, j > 1$, $R_i \cdot R_j \neq 0$ if and only if $|j - i| \leq 1$, and

$$\lim_{n\to\infty} \delta[p + \sum_n^\infty R_i] = 0.$$

† For a definition of this term, see § 6 below.

(6.1). LEMMA. *If p is any point of a space M and a is any point of $M-p$, there exists a simply infinite chain of regions from a converging to p.*

Proof. Let N_1 denote the component of $M-p$ containing a. Let $d_1 = 1/2\,\rho(p,a)$, and let x_1 be a point of the set $U_{d_1}(p)\cdot N_1$. Since p is a non-cut point of N_1+p, there exists a region R_1 such that $a+x_1 \subset R_1 \subset N_1$ and, $p\cdot \bar{R}_1 = 0$. Let $d_2 = 1/2\,\rho(p,\bar{R}_1)$, and let x_2 be a point of $U_{d_2}(p)\cdot N_2$, where N_2 is the component of $U_{d_1}(p)\cdot N_1$ containing x_1. Likewise there exists a region R_2 such that $x_1+x_2 \subset R_2 \subset N_2$ and $p\cdot\bar{R}_2 = 0$. Let $d_3 = 1/2\,\rho(p,\bar{R}_2)$ and let x_3 be a point of $U_{d_3}(p)\cdot N_3$, where N_3 is the component of $U_{d_2}(p)$ containing x_2. There exists a region R_3 such that $x_2+x_3 \subset R_3 \subset N_3$ and $p\cdot\bar{R}_3 = 0$, and so on. Continuing this process indefinitely, it is clear that the set $C = R_1+R_2+\cdots$ is the desired chain of regions from a converging to p. For, $\delta(p+\sum_n^\infty R_i) \leq \delta(p+N_n) \leq \delta[U_{d_n}(p)] = 2d_n \leq d_1/2^{n-1}$, and if $j-i>1$, $N_j \supset R_j$ but $N_j \subset U_{d_{j-1}}(p)$, therefore $N_j\cdot R_i = 0$.

(6.2) THEOREM. *If the point p of the space M is not an end point of M, then there exist two regions R_1 and R_2 such that $\bar{R}_1\cdot\bar{R}_2 = p$.*

Proof. If p is not a local end point, the theorem is obvious in view of proposition (5.2). Thus we may suppose p is a local end point but is neither an end point nor a cut point of M. Hence there exists a nodule N of M containing p, and p is a local end point of N. We now consider N as the space.

First we treat the case where N is locally divisible in p. There exists an $\epsilon>0$ such that p is an end point of the component U_0 of $V_\epsilon(p)$ containing p. There exists a region U such that p is an end point of U and the exterior and complement of U are connected and $\bar{U}\subset U_0$. There exists a region R_1 which is a subset of U and is the sum of a simply infinite chain of regions $C_1, C_2, \cdots$ converging to p, so that R_1+p is locally connected at p, and also so that $\bar{R}_1\subset U$. For each i, let x_i be a point of $C_i\cdot C_{i+1}$ and let $B_i = \sum_i^\infty C_j$. There exists a monotone decreasing sequence $G_1, G_2, \cdots$ of regions, subsets of U_0, such that $p = \prod_1^\infty \bar{G}_i\cdot U_0$, and for each i, $F(G_i)\cdot U_0 = p_i$, a single point. There exists an n_1 such that $G_{n_1}+p_{n_1}$ does not contain x_1. Since $F(G_{n_1})\cdot F(U_0)$ is necessarily $\neq 0$, for p_{n_1} cannot cut N, it follows that $G_{n_1}\cdot F(U)\neq 0$. Let y_1 be a point of $G_{n_1}\cdot F(U)$. There exists an $n_2>n_1$ such that $G_{n_2}+p_{n_2}$ does not contain y_1 or x_2. Let $E_0 = N-\bar{U}$, let K_1 be the component of $G_{n_1}-G_{n_1}\cdot S_1(\bar{B}_1)$ containing y_1, where we suppose our unit 1 so chosen that $\rho(y_1,\bar{B}_1)>1$; and let $E_1 = E_0+K_1$. Just as above

it follows that $G_{n_2} \cdot F(U) \supset$ some point y_2. There exists an $n_3 > n_2$ such that $G_{n_3} + p_{n_3}$ does not contain y_2 or x_3. Let K_2 be the component of $G_{n_2} - G_{n_2} \cdot S_{1/2}(\bar{B}_2)$ containing y_2 and let $E_2 = E_1 + K_2$. Let y_3 be a point of $G_{n_3} \cdot F(U)$. There exists an $n_4 > n_3$ such that G_{n_4} does not contain y_3 or x_4. Let K_3 be the component of $G_{n_3} - G_{n_3} \cdot S_{1/3}(\bar{B}_3)$ containing y_3 and let $E_3 = E_2 + K_3$, and so on. Continue this process indefinitely. Then since for each i, $p_{n_i} \subset B_{i+1}$ (for p_{n_i} separates x_i and p in U_0), it follows that for each i, $K_i \cdot G_{n_{i+1}} = 0$; and hence $K_i \cdot K_j = 0$ for $i \neq j$, and also $\bar{K}_i \cdot S_{1/i}(\bar{B}_i) \neq 0$. Hence if $R_2 = \sum_1^\infty E_n = E_0 + \sum_1^\infty K_i$, it is seen at once that R_2 is a region (since E_0 is connected) and that $\bar{R}_2 \supset p$ but $R_2 \cdot R_1 = 0$. It remains to show that $\bar{R}_1 \cdot \bar{R}_2 = p$. Now for each i, $F(E_i) \cdot \bar{R}_1 = 0$ and hence $\bar{R}_1 \cdot \sum_1^\infty F(E_i) = 0$. Now

$$F(R_2) \subset \sum_1^\infty F(E_i) + \prod_1^\infty \bar{G}_{n_i}, \quad \text{for} \quad E_i - E_{i-1} \subset K_i \subset G_{n_i}.$$

Hence

$$\bar{R}_1 \cdot F(R_2) \subset \bar{R}_1 \cdot \sum_1^\infty F(E_i) + \bar{R}_1 \cdot \prod_1^\infty \bar{G}_{n_i} \subset 0 + U \cdot \prod_1^\infty \bar{G}_{n_i} = p,$$

and therefore $\bar{R}_1 \cdot \bar{R}_2 = p$, for $R_1 \cdot R_2 = 0$.

Now in case N is not locally divisible in p, then by Theorem (4.2), there exists a biunivalued and continuous transformation T of N into a set N^* which is locally divisible in $p^* = T(p)$, and furthermore such that every region in N^* whose complement and exterior are connected is the image of a region in N. Now by virtue of the case treated above, there exist two regions R^*_1 and R^*_2 in N^* such that $\bar{R}^*_1 \cdot \bar{R}^*_2 = p^*$, for clearly p^* is not an end point of N^*. These regions may be chosen so that their complements and exteriors are connected. Set $R_1 = T^{-1}(R^*_1)$, $R_2 = T^{-1}(R^*_2)$. Then R_1 and R_2 are regions in N and indeed $\bar{R}_1 \cdot \bar{R}_2 = p$. For obviously $\bar{R}_1 \cdot \bar{R}_2 \subset p$, since T is biunivalued and continuous; and then if $\bar{R}_1 \cdot \bar{R}_2 \neq p$, it must be true that for either R_1 or R_2, say R_1, $\bar{R}_1 \cdot p = 0$; and hence by corollary (3.1a) there exists a region U in N containing p but such that $U \cdot \bar{R}_1 = 0$ and such that $U^* = T(U)$ is a region in N^*. Clearly this is impossible, for then $U^* \cdot R^*_1 = 0$, contrary to the fact that p^* is a limit point of R^*_1. Therefore $\bar{R}_1 \cdot \bar{R}_2 = p$, and our theorem is proved.

COROLLARY (6.2a). *In case M is locally divisible in the non-end point p, one of the sets R_1 and R_2, say R_1, may be chosen so that $R_1 + p$ is locally connected.*

COROLLARY (6.2b). *In order that the point p of a space M be an end point of M, either of the following conditions is necesary and sufficient:*

1). *that there should not exist two regions R_1 and R_2 such that $\bar{R}_1 \cdot \bar{R}_2 = p$.*

2). *that p should not be a cut point of any connected subset of M.*

7. *Generalized three-point theorem.* The theorem of Ayres † that any three points $p_1, p_2, p_3, \cdots$ of a cyclicly connected continuous curve C taken in any order p_i, p_j, p_k lie on an arc $p_i p_j p_k$ in C will be called *the three-point theorem.* We now prove an analogus theorem for spaces M.

(7.1). THEOREM. *If p is not a local end point of a set M having no cut point, then if a and b are any two distinct points of $M - p$, there exist regions R_a and R_b containing a and b respectively, and such that $\bar{R}_a \cdot \bar{R}_b = p$ and $R_a + p$ and $R_b + p$ are locally connected.*

Proof. In case p is not a local separating point of M, this theorem has been proved in the proof of Theorem (5.2). Hence we may suppose p is a local separating point of M. Thus there exists a region K containing p but neither a nor b and two points c and d which belong to different components K_c and K_d, respectively, of $K - p$. There exists a region L such that $\bar{L} \subset K$ and such that the points a, b, c, and d all lie together in a single component C^* of $M - \bar{L}$. Now there exist components L_c and L_d of $L - p$ which are subsets respectively of K_c and K_d. By virtue of (5.1) there exist two strongly separated regions Q_a and Q_b joining a and $L_c + L_d$, and b and $L_c + L_d$ respectively. Let $W = \overline{L_c + L_d + p}$ and let N_a, N_b and C denote the components of $Q_a - Q_a \cdot W$, $Q_b - Q_b \cdot W$ and $M - W$ respectively which contain a, b, and C^*, respectively. Now W contains at least one limit point of each of the sets N_a and N_b. And if for one of these sets, say N_a, p is the only limit point of N_a in W, then set $R_a = N_a$ and define R_b as follows: either $\bar{L}_c$ or $\bar{L}_d$, say $\bar{L}_c$, contains a point $q \neq p$ which is a limit point of N_b; let U_q be a region containing q and such that $\bar{U}_q \cdot \bar{N}_a = 0$; then set $R_b = N_b + U_q + L_c$. In this case clearly the sets R_a and R_b thus defined have the desired properties for the theorem. If this is not the case, then W contains points x and y, distinct from p, which are limit points of N_a and N_b respectively. Now if there is any choice of the points x and y such that x belongs to one of the sets $\bar{L}_c$ and $\bar{L}_d$ and y to the other one, say $x \subset \bar{L}_c$ and $y \subset \bar{L}_d$, then we define the sets R_a and R_b as follows: let U_x and U_y be strongly separated regions containing x and y respectively and such that

$$\bar{U}_x \cdot (\bar{N}_b + \bar{L}_d) = \bar{U}_y \cdot (\bar{N}_a + \bar{L}_c) = 0.$$

† See *Bulletin de l'Academie Polonaise des Sciences et des Lettres*, 1928, pp. 127-142.

Then $N_a + U_x$ and $N_b + U_y$ are mutually exclusive regions joining a and x, and b and y, respectively; clearly these regions contain strongly separated regions E_a and E_b joining a and x, and b and y, respectively. Then if $R_a = E_a + L_c$ and $R_b = E_b + L_d$, the sets R_a and R_b have the desired properties.

Now if no such choice of x and y is possible, then both x and y belong to one of the sets $\bar{L}_c$ and $\bar{L}_d$, say $\bar{L}_c$, and no point of $\bar{L}_d - p$ is a limit point of $N_a + N_b$. Now since $d \subset C^*$, it follows that at least one point z of $\bar{L}_d - p$ is a limit point of C. Let U_z be a region containing z and such that $U_z \cdot (\bar{N}_a + \bar{N}_b + \bar{L}_c) = 0$. Let f be a point of $U_z \cdot C$ and let C_f be the component of $C - C \cdot (\bar{N}_a + \bar{N}_b)$ containing f. Either $\bar{N}_a \cdot C$ or $\bar{N}_b \cdot C$, say $\bar{N}_a \cdot C$, (since the two cases obviously are alike) contains a point $g \neq p$ which is a limit point of C_f. Let U_g be a region containing g and such that $\bar{U}_g \cdot \bar{N}_b + \bar{L}_c) = 0$. Then U_g contains a point h of C_f, and C_f contains a region V_{fh} joining f and h and such that $\bar{V}_{fh} \subset C_f$ and hence so that $\bar{V}_{fh} \cdot (\bar{N}_b + \bar{L}_c) = 0$. Now let U_y be a region containing y and such that

$$U_y \cdot (\bar{N}_a + \bar{U}_g + \bar{V}_{fh} + \bar{U}_z + \bar{L}_g) = 0.$$

Then $\quad G_a = N_a + U_g + V_{fh} + U_z + L_d \quad$ and $\quad G_b = N_b + U_y + L_c$

are mutually exclusive regions joining a and L_d, and b and L_c, respectively, such that $G_a \cdot L_c = G_b \cdot L_d = 0$. The regions G_a and G_b contain regions V_a and V_b, respectively, joining a and L_d, and b and L_c, respectively, and such that $\bar{V}_a \subset G_a$ and $\bar{V}_b \subset G_b$. Hence if $R_a = V_a + L_d$ and $R_b = V_b + L_c$, then $\bar{R}_a \cdot \bar{R}_b = p$ and $R_a + p$ and $R_b + p$ are locally connected sets. Thus our theorem is established.

It is clear that essentially the same argument may be applied to prove the following slightly more general theorem.

(7.2). THEOREM. *If p is a non-local end point of a space M having no cut point and which is imbedded in a space M_0, if A and B are mutually exclusive closed subsets of M such that A does not separate p and B and B does not separate p and A in M, and if A_0 and B_0 are subsets of M_0 such that $A_0 \cdot M = A$ and $B_0 \cdot M = B$, then there exist regions R_1 and R_2 in M_0 such that $R_1 \cdot A \neq 0$, $R_2 \cdot B \neq 0$, $\bar{R}_1 \cdot B_0 = \bar{R}_2 \cdot A_0 = 0$, $\bar{R}_1 \cdot \bar{R}_2 = p$, and $R_1 + p$ and $R_2 + p$ are locally connected.*

8. GENERALIZED CYCLIC CONNECTIVITY THEOREM. *If a and b are any two non-local end points of a space M which has no cut point, then there exist two regions R_1 and R_2 such that $\bar{R}_1 \cdot \bar{R}_2 = a + b$ and $R_1 + a + b$ and $R_2 + a + b$ are locally connected sets.*

Proof. Now if $M - (a+b)$ is not connected, the theorem is obvious, for in this case we may take R_1 and R_2 as any two distinct components of $M - (a+b)$. Hence we may suppose that $M - (a+b) = (M-a) - b$ is connected. Now there exist two mutually exclusive regions G_1 and G_2 such that $\bar{G}_1 \cdot \bar{G}_2 = a$ and $\bar{G}_1 \cdot b = \bar{G}_2 \cdot b = 0$ and $G_1 + a$ and $G_2 + a$ are locally connected. Since $M - (a+b)$ is connected, it follows that there exists a region G_3 joining G_1 and G_2 and such that $\bar{G}_3 \cdot (a+b) = 0$. Since clearly a is neither a cut point nor an end point of the set $G_1 + G_2 + G_3 + a$, it therefore lies in a nodule N of $G_1 + G_2 + G_3 + a$, and $N \cdot b = 0$. Let x and y be any two points of N. Then by (7.1) there exist two regions R_x and R_y containing x and y respectively and such that $\bar{R}_x \cdot \bar{R}_y = b$ and $R_x + b$ and $R_y + b$ are locally connected sets. Let S_x and S_y respectively denote components of $R_x - R_x \cdot \bar{N}$ and $R_y - R_y \cdot \bar{N}$, respectively, such that $\bar{S}_x \cdot \bar{S}_y = b$ and $\bar{S}_x \cdot \bar{N} \neq 0 \neq \bar{S}_y \cdot \bar{N}$. Now it will be seen at once that all cases in which $\bar{S}_x \cdot N$ or $\bar{S}_y \cdot N$ are non-vacuous are only simpler than that in which both of these sets are vacuous. Hence we may assume that $\bar{S}_x \cdot N = \bar{S}_y \cdot N = 0$. Let x' and y' be points of $\bar{S}_x \cdot \bar{N}$ and $\bar{S}_y \cdot \bar{N}$ respectively, and let G_x and G_y be strongly separated regions containing x' and y' respectively and such that $\bar{G}_x \cdot (\bar{S}_y + a) = \bar{G}_y \cdot (\bar{S}_x + a) = 0$. Let N^* be the nodule of the set $N + G_x + G_y$ which contains N, let $A_0 = S_x$, $B_0 = S_y$, $A = N^* \cdot A_0$ and $B = N^* \cdot B_0$. Now A does not separate the sets B and a in N^*, for $G_y \cdot N^* + N$ is a connected subset of N^* which contains both a and the point y' of B but contains no point of A. Similarly, B does not separate the sets A and a in N^*. Thus since $A = A_0 \cdot N^*$ and $B = B_0 \cdot N^*$, it follows by Theorem (7.2) that there exist two regions T_x and T_y containing points of A and B respectively and such that $\bar{T}_x \cdot \bar{T}_y = a$, $\bar{T}_x \cdot B_0 = \bar{T}_y \cdot A_0 = 0$, and $T_x + a$ and $T_y + a$ are locally connected sets. Then if $R_1 = R_x + S_x$ and $R_2 = T_y + S_y$ it is seen at once that $\bar{R}_1 \cdot \bar{R}_2 = a + b$ and $R_1 + a + b$ and $R_2 + a + b$ are locally connected sets. This completes the proof of the theorem.

The theorem just proved is a generalization of the theorem,† which we call the *cyclic connectivity theorem,* that any locally compact space M which has no cut point is cyclicly connected. For clearly such a space M can have no local end point. Hence for any two points a and b, the regions R_1 and R_2 exist as in our theorem. Then $R_1 + a + b$ and $R_2 + a + b$ contain arcs

† See G. T. Whyburn, *Proceedings of the National Academy of Sciences,* Vol. 13 (1927), pp. 31-38; and W. L. Ayres, *American Journal of Mathematics,* Vol. 51 (1929), p. 590. For a simple proof of this theorem, see the author's paper, "On the Cyclic Connectivity Theorem", appearing in the *Bulletin of the American Mathematical Society.*

$(ab)_1$ and $(ab)_2$, respectively, from a to b; and obviously $(ab)_1 + (ab)_2$ is a simple closed curve in M.

COROLLARY (8a). *If the space M has no local separating point, then for each pair of points a and b of M, there exists an infinite sequence $R_1, R_2, R_3, \cdots$ of mutually exclusive regions such that for each i, $R_i + a + b$ is connected and locally connected and for each i and j, $i \neq j$, $\bar{R}_i \cdot \bar{R}_j = a + b$.*

Proof. Clearly M can have neither cut points nor local end points. Hence by the theorem in this section there exist two mutually exclusive regions R_1 and S_1 such that $\bar{R}_1 \cdot \bar{S}_1 = a + b$ and $R_1 + a + b$ and $S_1 + a + b$ are locally connected. Now S_1 can have no local separating points, and therefore it follows that a and b are not local end points of $S_1 + a + b$. Hence, by the theorem just established, there exist in S_1 two regions R_2 and S_2 such that $\bar{R}_2 \cdot \bar{S}_2 = a + b$ and $R_2 + a + b$ and $S_2 + a + b$ are locally connected. Similarly, S_2 can have no local separating point and there exist in S_2 two regions R_3 and S_3 such that $\bar{R}_3 \cdot \bar{S}_3 = a + b$ and $R_3 + a + b$ and $S_3 + a + b$ are locally connected, and so on. Continuing this process indefinitely, obviously the sets $R_1, R_2, R_3, \cdots$ so obtained have all the desired properties.

9. *Higher connectivity of absolute G_δ M-spaces.*

(9.1) THEOREM. *If the space M has no cut point and is an absolute G_δ set, then every two non-local end points a and b of M lie on a simple closed curve in M. Thus every nodular G_δ-space M which has no local end points is cyclicly connected.*

For by the theorem in § 8, there exist regions R_1 and R_2 in M such that $\bar{R}_1 \cdot \bar{R}_2 = a + b$ and $R_1 + a + b$ and $R_2 + a + b$ are locally connected; and since each of the sets $R_1 + a + b$ and $R_2 + a + b$ is also a G_δ space M, these sets contain † arcs $(ab)_1$ and $(ab)_2$ respectively from a to b. Clearly $(ab)_1 + (ab)_2$ is a simple closed curve in M containing $a + b$.

It should be noted that this theorem is also a generalization of the cyclic connectivity theorem.

(9.2). THEOREM. *If the G_δ-space M has no local separating points, then each pair of points a and b of M can be joined in M by a continuum T which is the sum of a set of $\mathbf{c}$ independent arcs from a to b, i. e., $T = \sum_{0 \leq x \leq 1} axb$, where axb is an arc in M from a to b and if $x \neq y$, $axb \cdot ayb = a + b$.*

This theorem may be established with the aid of the results in the pre-

† See the reference to Moore and Menger in § 2.

ceding sections by the same method as used in my paper, *Continuous Curves Without Local Separating Points*,† to prove the same theorem for a compact space M. The argument required here differs only in some of the details which must be modified slightly on account of the greater generality of the space, but is essentially the same proof.

10. *Conclusion. Some unsolved problems.* The proofs for a number of the propositions in the preceding sections are considerably complicated by our lack of knowledge of the existence of a property of spaces M analogus to the property of continuous curves C to the effect that each subcontinuum K of C is, for each $\epsilon > 0$, contained in a sub-continuous curve K^* of C such that $\delta(K^*) < \delta(K) + \epsilon$. This suggests the following problem, a positive solution to which would make possible greatly simplified proofs for some of the propositions in the present paper.

(10.1). PROBLEM. *If K is any connected subset of a space M, then is it true that for each $\epsilon > 0$, K is contained in a closed, connected and locally connected subset K^* of M such that $\delta(K^*) < \delta(K) + \epsilon$?*

It seems likely that a space M without cut points may possess a certain higher connectivity between any two of its points, even though these points be local end points. In this connection a solution to the following problem would be of interest.

(10.2). PROBLEM. *If a and b are any two points of a space M having no cut point, then does there exist two mutually exclusive regions R_1 and R_2 in M such that $\bar{R}_1 \cdot \bar{R}_2 \supset a + b$?* ‡

It may be noted here that it is not true that such regions R_1 and R_2 always exist so that $\bar{R}_1 \cdot \bar{R}_2 = a + b$. A simple modification of the space $\mathfrak{M}$ in § 3 shows that this is not always possible.

Finally, the author wishes to mention the desirability of a thorough investigation of the properties of the local end points of a space M and of the subsets of the set E of all local end points of M. For example, *can M be disconnected by the omission of E or of any closed subset of E?*

THE JOHNS HOPKINS UNIVERSITY.

---

† See *American Journal of Mathematics*, Vol. 53 (1931), pp. 163-166.

‡ Essentially this same problem was raised by the author in *Transactions of the American Mathematical Society*, Vol. 32 (1930), p. 943, problem (9.1). It is remarked that the results in the present paper greatly clarify the difficulties discussed in this reference.

GENERALIZED GREEN'S MATRICES FOR COMPATIBLE SYSTEMS OF DIFFERENTIAL EQUATIONS.*

By William T. Reid.†

1. *Introduction.* The existence and properties of the Green's matrix for an incompatible system of linear differential equations of the first order has been treated by Bounitzky, Birkhoff and Langer, Bliss and others.‡ In the systems considered by the above mentioned writers the coefficients of the system are supposed to be continuous functions of the independent variable. W. M. Whyburn § has demonstrated the existence of the Green's matrix for a system of equations whose coefficients are only Lebesgue summable functions of the independent variable.

For certain special types of compatible differential systems, where the differential equation is a single linear equation with continuous coefficients, the existence of generalized Green's functions has been shown by Hilbert, Bounitzky, Westfall and others.¶ Recently Elliott ‖ has treated generalized Green's functions for general compatible differential systems consisting of a single differential equation of the n-th order with continuous coefficients, together with boundary conditions involving the values of the solution and its first $n-1$ derivatives at two points.

It is the purpose of this paper to show that theorems corresponding to those obtained by Elliott are true for a system of n ordinary linear differential equations of the first order with Lebesgue summable coefficients, together with boundary conditions involving the values of the solution at two points.

Vector and matrix notation is used throughout the paper. A capital

* Presented to the American Mathematical Society, December 30, 1929.

† National Research Fellow in Mathematics.

‡ E. Bounitzky, *Journal de Mathématiques*, (6), Vol. 5 (1909), p. 65; G. D. Birkhoff and R. E. Langer, *Proceedings of the American Academy of Arts and Sciences*, Vol. 58 (1923), p. 51; G. A. Bliss, *Transactions of the American Mathematical Society*, Vol. 28 (1926), p. 561.

§ W. M. Whyburn, *Annals of Mathematics*, (2), Vol. 28 (1927), p. 291.

¶ D. Hilbert, *Göttinger Nachrichten*, (1904), p. 213, and *Grundzüge einer allgemeinen Theorie der Integralgleichungen*, Berlin, (1912), p. 44; E. Bounitzky, *loc. cit.*; W. Westfall, *Dissertation*, Göttingen, (1905), and *Annals of Mathematics*, (2), Vol. 10 (1909), p. 177.

‖ W. W. Elliott, *American Journal of Mathematics*, Vol. 50 (1928), p. 243; also, *American Journal of Mathematics*, Vol. 51 (1929), p. 397.

letter will denote a square matrix with n rows and n columns whose element in the i-th row and j-th column is given by the same letter with the subscript ij. E is used to denote the unit matrix, that is, $E_{ij}=0$ if $i \neq j$, $E_{ii}=1$, and 0 is the matrix each of whose elements is zero. It may be mentioned here that in this paper the idea of the zero matrix and the unit matrix is used only in connection with matrices whose elements are constants or continuous functions. If A and B are two square matrices, then the matrix product AB is the matrix C, where $C_{ij}=A_{i\alpha}B_{\alpha j}$.* Similarly, all vectors are supposed to have n components and if η is a vector, then η_i is the i-th component of η. The product $A\eta$ of a matrix A by the vector η is given by the vector b, where $b_i=A_{i\alpha}\eta_\alpha$ $(i=1,2,\cdots,n)$. Similarly, the product ηA is given by the vector c, where $c_i=\eta_\alpha A_{\alpha i}$ $(i=1,2,\cdots,n)$. The conjugate imaginary of a complex quantity a is denoted by $\bar{a}$, and the conjugate of the matrix A, which will be denoted by $\bar{A}$, is the matrix each element of which is the conjugate imaginary of the corresponding element of A. $\tilde{A} \equiv \| \tilde{A}_{ij} \|$ is used to denote the adjoint, or transpose, of A, i. e., $\tilde{A}_{ij}=A_{ji}$ $(i,j=1,2,\cdots,n)$.

A solution of the vector differential equation

$$(1) \qquad y'=A(x)y+g(x),$$

where $A_{ij}(x)$ $(i,j=1,2,\cdots,n)$ and $g_i(x)$ $(i=1,2,\cdots,n)$ are Lebesgue summable functions, real or complex, of the real variable x on the interval $X: a \leqq x \leqq b$, we define as an absolutely continuous vector $y(x) \equiv [y_\alpha(x)]$ which satisfies (1) on X_0.† With the homogeneous vector equation

$$(2) \qquad y'=A(x)y$$

we associate boundary conditions

$$(3) \qquad My(a)+Ny(b)=0,$$

where M and N are square matrices such that the matrix $\| M, N \|$ ‡ is of rank n. The differential equation adjoint to (2) is

$$(4) \qquad z'=-zA(x)$$

---

* The repetition of a subscript in an expression will denote summation with respect to that subscript over the values from 1 to n.

† X_0 is used to denote 'almost everywhere' on X. The excepted null set is not constant and may vary. Throughout this paper primes are used to denote differentiation with respect to an independent variable x.

‡ If M and N are two matrices of n rows and n columns, then $\| M; N \|$ is used to denote the matrix U which has n rows and $2n$ columns and whose elements are defined as: $U_{i,2j-1}=M_{ij}$, $U_{i,2j}=N_{ij}$ $(i,j=1,2,\cdots,n)$.

and the adjoint boundary conditions are given by

(5) $$z(a)P + z(b)Q = 0,$$

where the matrix of coefficients of (5) is of rank n and $MP - NQ = 0$.*

A matrix $G(x;t)$ is said to be a Green's matrix for the system (2), (3) if for each value of t on $a < t < b$ we have:

(6) Each column of $G(x;t)$ as a function of x is a solution of (2) on $a \leqq x < t$ and $t < x \leqq b$,

(7) $$G(t+;t) - G(t-;t) = E,$$

(8) $$MG(a;t) + NG(b;t) = 0.$$

Whenever the system (2), (3) is incompatible there exists a unique Green's matrix for the system which is given by

(9) $$G(x:t) = (1/2)Y(x)[\,|x-t|/(x-t)E + D\Delta]Z(t),$$

where $Y(x)$ and $Z(x)$ are matrix solutions † of (2) and (4) respectively such that $Y(x)Z(x) = Z(x)Y(x) = E$, D is the reciprocal of the matrix $MY(a) + NY(b)$, and $\Delta = MY(a) - NY(b)$. We have

(A) When the system (2), (3) is incompatible the unique solution of (1), (3), where $g(x) \equiv [g_a(x)]$ is any summable vector, is given by ‡

$$y(x) = \int_a^b G(x;t)g(t)\,dt.$$

(B) The functions $H_{ij}(x;t) = -G_{ji}(t;x)$ defined by (9) are the elements of the Green's matrix for the adjoint system (4), (5) and

(10) Each row of $G(x;t)$ as a function of t is a solution of (4) on $a \leqq t < x$ and $x < t \leqq b$,

(11) $$G(x;x-) - G(x;x+) = E,$$

(12) $$G(x;a)P + G(x;b)Q = 0.$$§

2. *Existence of Generalized Green's Matrices.* We shall now assume that the system (2), (3) is compatible of index r. Then the adjoint system (4),

---

* G. A. Bliss, *loc. cit.*, p. 564.

† A square matrix $Y(x)$ each column of which is a solution of (2) and such that the determinant $|Y(x)|$ is different from zero on X is called a matrix solution of (2). Similarly, a matrix $Z(x)$ each row of which is a solution of (4) and $|Z(x)| \neq 0$ on X is a matrix solution of (4).

‡ G. A. Bliss, *loc. cit.*, p. 579; also W. M. Whyburn, *loc. cit.* All integrals considered in this paper are taken in the sense of Lebesgue.

§ G. A. Bliss, *loc. cit.*, p. 578; also W. M. Whyburn, *loc. cit.*

(5) is also compatible of index r.* Let the matrix ${}^rY(x) \equiv \| {}^rY_{ij}(x) \|$ be defined as follows: ${}^rY_1, {}^rY_2, \cdots, {}^rY_r$ $[{}^rY_j \equiv ({}^rY_{aj})$ $(j = 1, 2, \cdots, r)]$ are r linearly independent solutions of the system (2), (3) and ${}^rY_{ij} = 0$ if $j > r$. Also by ${}^rZ(x)$ we denote a matrix whose first r rows are linearly independent solutions of the adjoint system (4), (5) and the elements of the remaining rows are all zero. The general solution of (2), (3) is then of the form

$$y(x) = \sum_{a=1}^{r} {}^rY_a(x) a_a,$$

where the a_a's are arbitrary constants. Whenever (2), (3) is compatible the system (1), (3) has a solution if and only if the vector equation

$$\int_a^b {}^rZ(x) g(x) dx = 0 \tag{13}$$

is satisfied by the vector $g(x)$.

If condition (13) is satisfied we seek to determine a matrix $G(x; t)$ which is continuous in (x, t) at every point on $a \leqq {x \atop t} \leqq b$ except along the line $x = t$, which satisfies the condition

$$G(t+; t) - G(t-; t) = E, \tag{14}$$

and is such that every solution of (1), (3) may be written in the form

$$y(x) = \int_a^b G(x; t) g(t) dt + \sum_{a=1}^{r} {}^rY_a(x) a_a. \tag{15}$$

If such a matrix $G(x; t)$ exists we will say that $G(x; t)$ is a *Generalized Green's Matrix* for the compatible system (2), (3).

Let $Y(x) \equiv \| Y_{ij}(x) \|$ be a matrix solution of (2) such that the first r columns of $Y(x)$ are ${}^rY_1, {}^rY_2, \cdots, {}^rY_r$. By $Z(x)$ we denote the matrix solution of (4) such that $Z(x)Y(x) = Y(x)Z(x) = E$ on X. Let $s_i(Y_j)$ denote $M_{ia}Y_{aj}(a) + N_{ia}Y_{aj}(b)$ $(i, j = 1, 2, \cdots, n)$. If the system (2), (3) has r linearly independent solutions, then the matrix

---

* For the case in which the elements of the matrix $A(x)$ are real continuous functions of the variable x and the elements of M and N are real constants this result has been established by Bliss and his method of proof carries over wholly for the system (2), (3). See G. A. Bliss, *loc. cit.*, pp. 566-567. In a recent paper the author has also established this result for an infinite system of linear differential equations which includes as a special case the finite system (2), (3) when the coefficients of the system are real. See W. T. Reid, *Transactions of the American Mathematical Society*, Vol. 32 (1930), pp. 284-318; in particular, pp. 306-311.

$$\left\| \begin{matrix} s_1(Y_{r+1}) & s_1(Y_{r+2}) & \cdot\ \cdot & s_1(Y_n) \\ s_2(Y_{r+1}) & \cdot & \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ s_n(Y_{r+1}) & \cdot & \cdot\ \cdot & s_n(Y_n) \end{matrix} \right\|$$

is of rank $n-r$. The integers $k_1, k_2, \cdots, k_{n-r}$ may then be so chosen that the matrix

$$\left\| \begin{matrix} s_{k_1}(Y_{r+1}) & \cdot\ \cdot & s_{k_1}(Y_n) \\ s_{k_2}(Y_{r+1}) & \cdot\ \cdot & \cdot \\ \cdot & \cdot\ \cdot & \cdot \\ s_{k_{n-r}}(Y_{r+1}) & \cdot\ \cdot & s_{k_{n-r}}(Y_n) \end{matrix} \right\|$$

has a unique reciprocal, which we will denote by

$$\left\| \begin{matrix} s^{-1}{}_{11} & s^{-1}{}_{12} & \cdot\ \cdot & s^{-1}{}_{1,n-r} \\ s^{-1}{}_{21} & \cdot & \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ s^{-1}{}_{n-r,1} & \cdot & \cdot\ \cdot & s^{-1}{}_{n-r,n-r} \end{matrix} \right\| .$$

Now define the matrix $\mathfrak{D} \equiv \| \mathfrak{D}_{ij} \|$ as: $\mathfrak{D}_{r+i,k_j} = s_{ij}{}^{-1}$ $(i, j = 1, 2, \cdots n-r)$; $\mathfrak{D}_{ij} = 0$ if $i \leqq r$ or $j \neq k_\alpha$ $(\alpha = 1, 2, \cdots, n-r)$.

THEOREM 1. *A generalized Green's matrix for the compatible system* (2), (3) *exists and may be written as*

(16) $$G(x;t) = (1/2)Y(x)[\,|x-t|/(x-t)E + \mathfrak{D}\Delta]Z(t),$$

where $Y(x)$, $Z(x)$ *and* $\mathfrak{D}$ *are defined as above and* $\Delta = MY(a) - NY(b)$.

Let $K(x;t) = (1/2)Y(x)Z(t)[\,|x-t|/(x-t)]$. Then $K(x;t)$ as a function of x is a matrix solution of (2) on $a \leqq x < t$ and $t < x \leqq b$. Furthermore,

$$K(t+;t) - K(t-;t) = E.$$

Let
$$u(x) = \int_a^b K(x;t)g(t)dt,$$
$$= (1/2)Y(x)[\int_a^x Z(t)g(t)dt - \int_x^b Z(t)g(t)dt].$$

Since $u(x)$ is clearly absolutely continuous on X and

$$u'(x) = A(x)u(x) + g(x)$$

on X_0, the vector $u(x)$ is a solution of (1). Then every solution of (1) is of the form

$$y(x) = u(x) + \sum_{\alpha=1}^{n} Y_\alpha(x)a_\alpha.$$

Since by hypothesis the system (1), (3) is compatible, there exist values of the a_α's so that

$$\sum_{\alpha=1}^{n} s_i(Y_\alpha)a_\alpha + s_i(u) = 0 \qquad (i = 1, 2, \cdots, n).$$

Since $s_i(Y_j) = 0$ $(i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, r)$, we have

$$\sum_{\alpha=1}^{n-r} s_{k_i}(Y_{r+\alpha})a_{r+\alpha} = -\, s_{k_i}(u) \qquad (i = 1, 2, \cdots, n-r),$$

$$= (1/2)\,[M_{k_i\alpha}Y_{\alpha\beta}(a) - N_{k_i\alpha}Y_{\alpha\beta}(b)] \int_a^b Z_{\beta\gamma}(t)g_\gamma(t)\,dt.$$

Then

$$a_{r+i} = (1/2) \sum_{\mu=1}^{n-r} \mathfrak{D}_{r+i,k_\mu}[M_{k_\mu\alpha}Y_{\alpha\beta}(a) - N_{k_\mu\alpha}Y_{\alpha\beta}(b)] \int_a^b Z_{\beta\gamma}(t)g_\gamma(t)\,dt \qquad (i = 1, 2, \cdots, n-r).$$

Therefore

$$y(x) = \sum_{\alpha=1}^{r} {}^rY_\alpha(x)a_\alpha + \int_a^b (1/2)Y(x)[\,|x-t|/(x-t)E + \mathfrak{D}\Delta]Z(t)g(t)\,dt,$$

and we have that a generalized Green's matrix for (2), (3) is given by (16).

It is to be noted that the choice of the integers $k_1, k_2, \cdots, k_{n-r}$ is in general not unique. When $r = 0$ we have from (16) the ordinary Green's matrix for the incompatible system as given by (9).

THEOREM 2. *The generalized Green's matrix for the compatible system* (2), (3) *is not unique. If* $G_1(x;t)$ *is one generalized Green's matrix, then every generalized Green's matrix is of the form*

$$G(x;t) = G_1(x;t) + {}^rY(x)U(t) + V(x)\,{}^rZ(t), \tag{17}$$

where $V(x)$ *and* $U(x)$ *are matrices each element of which is continuous on* X. *Furthermore, every function* $G(x;t)$ *of the form* (17) *is a generalized Green's matrix for the compatible system* (2), (3).

It is evident that there exist constant matrices T and T^* such that ${}^rY(x)T = \mathbf{Y}(x)$ and $T^{*\,r}Z(x) = \mathbf{Z}(x)$, where: (a) the first r columns of $\mathbf{Y}(x)$ and the first r rows of $\mathbf{Z}(x)$ are linearly independent solutions of (2), (3) and (4), (5) respectively; (b) $\mathbf{Y}_{ij}(x) = 0 = \mathbf{Z}_{ji}(x)$ if $j > r$;

(c) $$\int_a^b \sum_{\alpha=1}^{n} \bar{\mathbf{Y}}_{\alpha i}(x)\mathbf{Y}_{\alpha j}(x)\,dx = E_{ij} = \int_a^b \sum_{\alpha=1}^{n} \bar{\mathbf{Z}}_{i\alpha}(x)\mathbf{Z}_{j\alpha}(x)\,dx \qquad (i, j = 1, 2, \cdots, r);$$

(d) $$T_{ij}=0=T^*_{ij} \quad \text{if} \quad i>r \quad \text{or} \quad j>r.$$

Let

$$g_i(x)=\sum_{\alpha=1}^{r}\bar{\mathbf{Z}}_{\alpha i}(x)c_\alpha+h_i(x) \qquad (i=1,2,\cdots,n),$$

where $h(x)\equiv[h_\alpha(x)]$ is any summable vector on X and

$$c_i=-\int_a^b\sum_{\alpha=1}^{n}\mathbf{Z}_{i\alpha}(v)h_\alpha(v)dv \qquad (i=1,2,\cdots,r).$$

In view of (13) we have that for every vector $g(x)\equiv[g_\alpha(x)]$ so determined the system (1), (3) is compatible. If $G_1(x;t)$ and $G(x;t)$ are two generalized Green's matrices for (2), (3), let $G(x;t)-G_1(x;t)=D(x;t)$. $D(x;t)$ is continuous in (x,t) on $a\leqq {x \atop t}\leqq b$. Then

$$\begin{aligned}(18)\qquad &\int_a^b D_{i\beta}(x;t)g_\beta(t)dt\\ &=\int_a^b D_{i\beta}(x;t)[-\sum_{\alpha=1}^{r}\bar{\mathbf{Z}}_{\alpha\beta}(t)\{\int_a^b\mathbf{Z}_{\alpha\gamma}(v)h_\gamma(v)dv\}+h_\beta(t)]dt,\\ (19)\qquad &=\int_a^b[-\sum_{\alpha=1}^{r}\{\int_a^b D_{i\beta}(x;v)\mathbf{Z}_{\alpha\beta}(v)dv\}\bar{\mathbf{Z}}_{\alpha\gamma}(t)+D_{i\gamma}(x;t)]h_\gamma(t)dt.\dagger\end{aligned}$$

We also have

$$(20)\qquad \int_a^b D_{i\beta}(x;t)g_\beta(t)dt=\sum_{\mu=1}^{r}\mathbf{Y}_{i\mu}(x)d_\mu,$$

where the d_μ's are constants. Now

$$\begin{aligned}d_\mu&=\int_a^b\bar{\mathbf{Y}}_{\sigma\mu}(u)\left(\sum_{\gamma=1}^{r}\mathbf{Y}_{\sigma\gamma}(u)d_\gamma\right)du \qquad (\mu=1,2,\cdots,r),\\ (21)\qquad &=\int_a^b\bar{\mathbf{Y}}_{\sigma\mu}(u)\left(\int_a^b[-\sum_{\alpha=1}^{r}\left\{\int_a^b D_{\sigma\beta}(u;v)\bar{\mathbf{Z}}_{\alpha\beta}(v)dv\right\}\mathbf{Z}_{\alpha\gamma}(t)\right.\\ &\qquad\left.+D_{\sigma\gamma}(u;t)]h_\gamma(t)dt\right)du\end{aligned}$$

in view of (19) and (20). From (19), (20) and (21) we then obtain

$$\begin{aligned}&\int_a^b\Big[-\sum_{\alpha=1}^{r}\left\{\int_a^b D_{i\beta}(x;v)\bar{\mathbf{Z}}_{\alpha\beta}(v)dv\right\}\mathbf{Z}_{\alpha\gamma}(t)+D_{i\gamma}(x;t)\\ (22)\qquad &+\sum_{\mu=1}^{r}\mathbf{Y}_{i\mu}(x)\left(\sum_{\alpha=1}^{r}\Big[\int_a^b\bar{\mathbf{Y}}_{\sigma\mu}(u)\left\{\int_a^b D_{\sigma\beta}(u;v)\bar{\mathbf{Z}}_{\alpha\beta}(v)dv\right\}du\Big]\mathbf{Z}_{\alpha\gamma}(t)\right.\\ &\qquad\left.-\int_a^b\bar{\mathbf{Y}}_{\sigma\mu}(u)D_{\sigma\gamma}(u;t)du\right)\Big]h_\gamma(t)dt=0.\end{aligned}$$

† The change of the order of integration in the double integral occurring in (18) is clearly permissible since the integrand is the sum of r terms, each of which is the product of a continuous function of t and a summable function of v. This also follows from a general theorem due to Carathéodory. See C. Carathéodory, *Vorlesungen ueber reelle Funktionen*, Leipzig, 1918, p. 641.

The change of the order of integration necessary to obtain relation (22) is, as before, permissible. Since the coefficient of $h_\gamma(t)$ $(\gamma = 1, 2, \cdots, n)$ in (22) is a continuous function on $a \leqq t \leqq b$ for each value of x on X, and the relation (22) is true for every summable vector $h(x)$, we have

$$(23) \qquad D_{i\gamma}(x;t) = \sum_{\alpha=1}^{r} V^*_{i\alpha}(x)\mathbf{Z}_{\alpha\gamma}(t) + \sum_{\beta=1}^{r} \mathbf{Y}_{i\beta}(x)U^*_{\beta\gamma}(t) \qquad (i, \gamma = 1, 2, \cdots, n),$$

where

$$V^*_{ij}(x) = \int_a^b D_{i\beta}(x;v)\bar{\mathbf{Z}}_{j\beta}(v)dv \qquad (i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, r),$$

$$U^*_{ij}(t) = -\sum_{\alpha=1}^{r}\left[\int_a^b \bar{\mathbf{Y}}_{\sigma i}(u)\left\{\int_a^b D_{\sigma\beta}(u;v)\bar{\mathbf{Z}}_{\alpha\beta}(v)dv\right\}du\right]\mathbf{Z}_{\alpha j}(t)$$
$$+ \int_a^b \bar{\mathbf{Y}}_{\sigma i}(u)D_{\sigma j}(u;t)du \qquad (i = 1, 2, \cdots, r;\ j = 1, 2, \cdots, n).$$

In view of (14) we have that each element of $D(x;t)$ is continuous in (x, t) on $a \leqq {x \atop t} \leqq b$ and therefore, since each element of $\mathbf{Z}(x)$ and $\mathbf{Y}(x)$ is continuous on X, it follows that $V^*_{ij}(x)$ and $U^*_{ij}(t)$ are continuous functions of their arguments. Let

$$U_{ij}(x) = \sum_{\alpha=1}^{r} T_{i\alpha}U^*_{\alpha j}(x) \quad \text{and} \quad V_{ij}(x) = \sum_{\alpha=1}^{r} V^*_{i\alpha}(x)T^*_{\alpha j} \qquad (i, j = 1, 2, \cdots, n).$$

Then

$$G(x;t) = G_1(x;t) + {}^rY(x)U(t) + V(x){}^rZ(t)$$

and clearly $U(x)$ and $V(x)$ are continuous on X. Also if $U(x)$ and $V(x)$ are any two matrices of continuous elements we have in view of (13) that if $G_1(x;t)$ is a generalized Green's matrix for (2), (3), then $G_1(x;t) + {}^rY(x)U(t) + V(x){}^rZ(t)$ is also a generalized Green's matrix for the compatible system (2), (3).

3. *Principal Generalized Green's Matrices.* We will now determine a generalized Green's matrix which shall possess for adjoint systems the same property as is given for the ordinary Green's matrix by (B) in the introduction.

Let M^*, N^*, P^* and Q^* be square matrices such that the matrices

$$\begin{Vmatrix} M_{ij} & N_{ij} \\ M^*_{ij} & N^*_{ij} \end{Vmatrix} \qquad \begin{Vmatrix} -P^*_{ij} & -P_{ij} \\ Q^*_{ij} & Q_{ij} \end{Vmatrix}$$

are unique reciprocals. For any pair of vectors $y(x)$ and $z(x)$ defined on X the vectors $s(y)$, $s^*(y)$, $t(z)$ and $t^*(z)$ are defined as

$$
(24) \qquad \begin{aligned} s(y) &= My(a) + Ny(b), & t(z) &= z(a)P + z(b)Q, \\ s^*(y) &= M^*y(a) + N^*y(b), & t^*(z) &= z(a)P^* + z(b)Q^*. \end{aligned}
$$

Then we have the identity †

$$
(25) \qquad s_\alpha(y)t^*_\alpha(z) + s^*_\alpha(y)t_\alpha(z) = y_\alpha(x)z_\alpha(x) \Big|_{x=a}^{x=b}
$$

If $y(x)$ and $z(x)$ are solutions of (2) and (4) respectively, then

$$
[z_\alpha(x)y_\alpha(x)]' = z_\alpha(x)y_\alpha'(x) + z_\alpha'(x)y_\alpha(x) = 0
$$

on X_0, and therefore $z_\alpha(x)y_\alpha(x)$ is a constant for x on X. Let $z(x)$ be any solution of (4) and $K(x;t) = (1/2)Y(x)Z(t)\,|x-t|/(x-t)$, where $Y(x)$ and $Z(x)$ are matrix solutions of (2) and (4) respectively which are reciprocals for x on X. Then

$$
\begin{aligned}
0 = \int_a^b \{ z_\alpha(x)[(\partial/\partial x)K_{\alpha j}(x;t) - A_{\alpha\beta}(x)K_{\beta j}(x;t)] \\
+ [z_\alpha'(x) + z_\beta(x)A_{\beta\alpha}(x)]K_{\alpha j}(x;t)\}dx \\
(j = 1, 2, \cdots, n), \\
= z_\alpha(x)K_{\alpha j}(x;t) \Big|_{x=a}^{x=t-} + z_\alpha(x)K_{\alpha j}(x;t) \Big|_{x=t+}^{x=b}.
\end{aligned}
$$

Hence

$$
(26) \qquad \begin{aligned} z_\alpha(x)K_{\alpha j}(x;t) \Big|_{x=a}^{x=b} &= z_\alpha(x)K_{\alpha j}(x;t) \Big|_{x=t-}^{x=t+}, \\ &= z_j(t) \qquad (j = 1, 2, \cdots, n). \end{aligned}
$$

Let $K_j(x;t)$ denote the vector $[K_{\alpha j}(x;t)]$ which is the j-th column of $K(x;t)$. In view of relation (25) we have

LEMMA 1. *If $K(x;t) = (1/2)Y(x)Z(t)\,|x-t|/(x-t)$, where $Y(x)$ and $Z(x)$ are matrix solutions of* (2) *and* (4) *respectively which are unique reciprocals on X, then for every solution $z(x) \equiv [z_\alpha(x)]$ of* (4) *we have*

$$
t^*_\alpha(z)s_\alpha(K_j)_x + t_\alpha(z)s^*_\alpha(K_j)_x = z_j(t)\ddagger \qquad (j = 1, 2, \cdots, n).
$$

As before, let ${}^rZ(x)$ denote a matrix whose first r rows are linearly independent solutions of the system (4), (5) and the elements of the remaining $n-r$ rows are all zero. Let $\Psi(x)$ denote any matrix of Lebesgue summable functions of the form

$$
(27) \qquad \Psi(x) = \left\| \begin{matrix} \Psi_{11}(x) & \Psi_{12}(x) & \cdot & \cdot & \Psi_{1r}(x) & 0 & \cdot & 0 \\ \Psi_{21}(x) & \cdot & \cdot & \cdot & \Psi_{2r}(x) & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & 0 \\ \Psi_{n1}(x) & \cdot & \cdot & \cdot & \Psi_{nr}(x) & 0 & \cdot & 0 \end{matrix} \right\|,
$$

† See Bliss, *loc. cit.*, p. 565.

‡ The subscript x is used to denote that the vectors $s(K_j)$ and $s^*(K_j)$ are determined by (24) when K_j is considered as a function of x.

and such that the r-rowed determinant whose general element is

$$\int_a^b {}^rZ_{i\alpha}(x)\Psi_{\alpha j}(x)\,dx$$

is different from zero. Since the first r rows of ${}^rZ(x)$ are linearly independent solutions of (4), (5) it follows that there always exists a matrix $\Psi(x)$ satisfying the above condition. Let

$$R = \begin{Vmatrix} R_{11} & R_{12} & \cdot & R_{1r} & 0 & \cdot & 0 \\ R_{21} & \cdot & \cdot & R_{2r} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 & \cdot & 0 \\ R_{r1} & \cdot & \cdot & R_{rr} & 0 & \cdot & 0 \\ 0 & \cdot & \cdot & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & 0 & \cdot & 0 \end{Vmatrix},$$

where the r-rowed square matrix

$$\begin{Vmatrix} R_{11} & R_{12} & \cdot & \cdot & R_{1r} \\ R_{21} & \cdot & \cdot & \cdot & R_{2r} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ R_{r1} & \cdot & \cdot & \cdot & R_{rr} \end{Vmatrix}$$

is the unique reciprocal of the matrix

$$(28)\qquad \begin{Vmatrix} \int_a^b {}^rZ_{1\alpha}(x)\Psi_{\alpha 1}(x)\,dx & \cdot & \cdot & \int_a^b {}^rZ_{1\alpha}(x)\Psi_{\alpha r}(x)\,dx \\ \int_a^b {}^rZ_{2\alpha}(x)\Psi_{\alpha 1}(x)\,dx & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \int_a^b {}^rZ_{r\alpha}(x)\Psi_{\alpha 1}(x)\,dx & \cdot & \cdot & \int_a^b {}^rZ_{r\alpha}(x)\Psi_{\alpha r}(x)\,dx \end{Vmatrix}.$$

We now define the matrix $N(x;t)$ as

$$N(x;t) = -\Psi(x)R\,{}^rZ(t).$$

Then

$$\begin{aligned} \int_a^b {}^rZ_{i\alpha}(x)N_{\alpha j}(x;t)\,dx &= -\int_a^b {}^rZ_{i\alpha}(x)\Psi_{\alpha\beta}(x)R_{\beta\gamma}{}^rZ_{\gamma j}(t)\,dx, \\ &= -{}^rZ_{ij}(t), \\ &= -t^*_\alpha({}^rZ_i)s_\alpha(K_j)_x \\ &\qquad (i=1,2,\cdots,r;\ j=1,2,\cdots,n), \end{aligned}$$

in view of Lemma 1 if we denote by ${}^rZ_i$ the vector $[{}^rZ_{i\alpha}(x)]$.

Let $\mathfrak{Y}^*(x)$ be a matrix of absolutely continuous functions satisfying on X_0 the matrix differential equation

$$(d/dx)\mathfrak{Y}^*(x) = A(x)\mathfrak{Y}^*(x) - \Psi(x)R.$$

Then $\mathfrak{Y}(x) = \mathfrak{Y}^*(x)\,{}^rZ(t)$ is continuous in (x, t) and $\mathfrak{Y}_j(x;t) \equiv [\mathfrak{Y}_{aj}(x;t)]$ is a particular solution of the non-homogeneous equation

$$(29) \qquad (\partial/\partial x)\mathfrak{Y}_j = A(x)\mathfrak{Y}_j + N_j(x;t) \qquad (j = 1, 2, \cdots, n),$$

where $N_j(x;t)$ is the vector $[N_{aj}(x;t)]$. Then we have

$$\begin{aligned}\int_a^b {}^rZ_{ia}(x)N_{aj}(x;t)\,dx &= \int_a^b (\partial/\partial x)\,[{}^rZ_{ia}(x)\mathfrak{Y}_{aj}(x;t)]\,dx,\\ &= {}^rZ_{ia}(x)\mathfrak{Y}_{aj}(x;t)\,\Big|_{x=a}^{x=b},\\ &= t^*_a({}^rZ_i)s_a(\mathfrak{Y}_j)_x\\ &\qquad (i = 1, 2, \cdots, r;\ j = 1, 2, \cdots, n),\end{aligned}$$

in view of relation (25). Therefore we have

$$(30) \qquad t^*_a({}^rZ_i)s_a(\mathfrak{Y}_j)_x = -\,t^*_a({}^rZ_i)s_a(K_j)_x \qquad (i = 1, 2, \cdots, r;\ j = 1, 2, \cdots, n).$$

Let $Y(x) \equiv \| Y_{ij}(x) \|$ be a matrix solution of (2) and Y_j denote the vector (Y_{aj}). From (25) we have

$$t^*_a({}^rZ_i)s_a(Y_j) = 0 \qquad (i = 1, 2, \cdots, r;\ j = 1, 2, \cdots, n),$$

and furthermore the vectors $t^*({}^rZ_i)$ $(i = 1, 2, \cdots, r)$ are linearly independent.† Then in view of (30) we may determine quantities $C_{ij}(t)$ such that

$$s(Y_a)C_{aj} + s(\mathfrak{Y}_j)_x = -\,s(K_j)_x \qquad (i = 1, 2, \cdots, n).$$

Furthermore, since $s(\mathfrak{Y}_j)_x$ and $s(K_j)_x$ are continuous functions of t, the elements of the matrix $C(t) \equiv \| C_{ij}(t) \|$ may be chosen continuous functions of t. We have then established the following

LEMMA 2. *Each element of the matrix $\Upsilon(x;t) = \mathfrak{Y}(x;t) + Y(x)C(t)$ is continuous in (x, t) on $a \leqq {x \atop t} \leqq b$, is absolutely continuous in x on X for each fixed value of t on X, and is such that on X_0 the matrix differential equation*

$$(31) \qquad (\partial/\partial x)\Upsilon(x;t) = A(x)\Upsilon(x;t) + N(x;t)$$

† See Bliss, *loc. cit.*, p. 566.

is satisfied, and

$$s(\Upsilon_j)_x = -s(K_j)_x \qquad (j = 1, 2, \cdots, n). \tag{32}$$

The solution of (31), (32), is not unique but is determined except for added solutions of the homogeneous system (2), (3). Let $\Theta(x)$ be any matrix of Lebesgue summable functions of the form

$$\Theta(x) = \begin{Vmatrix} \Theta_{11}(x) & \Theta_{12}(x) & \cdot\ \cdot & \Theta_{1n}(x) \\ \Theta_{21}(x) & \cdot & \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ \Theta_{r1}(x) & \cdot & \cdot\ \cdot & \Theta_{rn}(x) \\ 0 & 0 & \cdot\ \cdot & 0 \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ 0 & 0 & \cdot\ \cdot & 0 \end{Vmatrix}.$$

and such that the r-rowed determinant whose general element is

$$\int_a^b \Theta_{i\alpha}(x)\,{}^rY_{\alpha j}(x)\,dx$$

is different from zero. Then there exists a matrix $L(t) \equiv \| L_{ij}(t) \|$ each element of which is continuous on $a \leqq t \leqq b$ and such that the matrix

$$G(x;t) = K(x;t) + \Upsilon(x;t) + {}^rY(x)L(t) \tag{33}$$

satisfies the relation

$$\int_a^b \Theta(x)G(x;t)\,dx = 0.$$

We have therefore established the following

THEOREM 3. *If the system* (2), (3) *is compatible of index* r, *then for each pair of matrices* $\Psi(x)$ *and* $\Theta(x)$ *defined as above there exists a matrix* $G(x;t)$ *which is continuous in* (x,t) *on* $a \leqq {}^x_t \leqq b$ *except at* $x = t$, *which is absolutely continuous in* x *on* $a \leqq x < t$ *and* $t < x \leqq b$, *and which is such that*

$$(\partial/\partial x)G(x;t) = A(x)G(x;t) + N(x;t) \quad on \ X_0, \tag{34}$$

$$G(t+;t) - G(t-;t) = E, \tag{35}$$

$$MG(a;t) + NG(b;t) = 0, \tag{36}$$

$$\int_a^b \Theta(x)G(x;t)\,dx = 0. \tag{37}$$

THEOREM 4. *The matrix* $G(x;t)$ *satisfying the conditions of Theorem* 3 *is a generalized Green's matrix for the compatible system* (2), (3).

Let

$$u(x) = \int_a^b G(x;t)g(t)\,dt,$$

where the form of $G(x;t)$ is given by (33). Then in view of (13),

$$u(x) = Y(x) \int_a^b [(1/2)Z(t) \mid x - t \mid /(x-t) + C(t)]g(t)\,dt + {}^rY(x) \int_a^b L(t)g(t)\,dt.$$

Then $u(x)$ is absolutely continuous on X and on X_0 we have

$$u'(x) = A(x)u(x) + g(x).$$

From (36) we obtain $Mu(a) + Nu(b) = 0$ and therefore $u(x)$ satisfies the compatible semi-homogeneous system (1), (3). Then every solution $y(x)$ of the system (1), (3) may be written in the form

$$y(x) = \sum_{\alpha=1}^{r} {}^rY_\alpha(x)a_\alpha + \int_a^b G(x;t)g(t)\,dt,$$

and therefore the matrix $G(x;t)$ defined by (33) is a generalized Green's matrix for (2), (3).

A generalized Green's matrix for (2), (3) which satisfies the conditions of Theorem 3 will be called a *principal generalized Green's matrix* for the compatible system (2), (3).

Corollary. *If* $u(x) = \int_a^b G(x;t)f(t)\,dt$, *where* $G(x;t)$ *is a principal generalized Green's matrix for the system* (2), (3) *and* $f(x) \equiv [f_\alpha(x)]$ *is any summable vector, then*

$$\begin{aligned} &u'(x) = A(x)u(x) + f(x) + \int_a^b N(x;t)f(t)\,dt \text{ on } X_0, \\ (38)\qquad &Mu(a) + Nu(b) = 0, \\ &\int_a^b \Theta(x)u(x)\,dx = 0. \end{aligned}$$

Now let

$$\mathfrak{N} = \begin{Vmatrix} \mathfrak{N}_{11} & \mathfrak{N}_{12} & \cdot & \cdot & \mathfrak{N}_{1r} & 0 & \cdot & 0 \\ \mathfrak{N}_{21} & \cdot & \cdot & \cdot & \mathfrak{N}_{1r} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & 0 \\ \mathfrak{N}_{r1} & \cdot & \cdot & \cdot & \mathfrak{N}_{rr} & 0 & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 & 0 & \cdot & 0 \end{Vmatrix},$$

where the r-rowed matrix

$$\begin{Vmatrix} \mathfrak{R}_{11} & \mathfrak{R}_{12} & \cdot & \cdot & \mathfrak{R}_{1r} \\ \mathfrak{R}_{21} & \cdot & \cdot & \cdot & \mathfrak{R}_{2r} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathfrak{R}_{r1} & \cdot & \cdot & \cdot & \mathfrak{R}_{rr} \end{Vmatrix}$$

is the unique reciprocal of

$$\begin{Vmatrix} \int_a^b \Theta_{1\alpha}(x)\,{}^rY_{\alpha 1}(x)\,dx & \cdot\ \cdot & \int_a^b \Theta_{1\alpha}(x)\,{}^rY_{\alpha r}(x)\,dx \\ \cdot & \cdot\ \cdot & \cdot \\ \cdot & \cdot\ \cdot & \cdot \\ \int_a^b \Theta_{r\alpha}(x)\,{}^rY_{\alpha 1}(x)\,dx & \cdot\ \cdot & \int_a^b \Theta_{r\alpha}(x)\,{}^rY_{\alpha r}(x)\,dx \end{Vmatrix}. \tag{39}$$

With respect to the matrices $\Psi(x)$ and $\Theta(x)$ a matrix $H(x;t)$ is said to be a *principal generalized Green's matrix* for the compatible adjoint system if each element is continuous in (x, t) on $a \leqq {x \atop t} \leqq b$ except at $x = t$, if each element is absolutely continuous in x on $a \leqq x < t$ and $t < x \leqq b$, and if furthermore:

$$(\partial/\partial x)\tilde{H}(x;t) = -\tilde{H}(x;t)A(x) - {}^rY(t)\mathfrak{R}\Theta(x) \text{ on } X_0, \tag{40}$$

$$H(t+;t) - H(t-;t) = E, \tag{41}$$

$$\tilde{H}(a;t)P + \tilde{H}(b;t)Q = 0, \tag{42}$$

$$\int_a^b \tilde{H}(x;t)\Psi(x)\,dx = 0. \tag{43}$$

The existence of a principal generalized Green's matrix for the adjoint system may be established by argument similar to that used in the proof of Theorem 3.

THEOREM 5. *If $G(x;t)$ and $H(x;t)$ are principal generalized Green's matrices for the system* (2), (3) *and the adjoint system* (4), (5) *respectively, with respect to a pair of matrices $\Psi(x)$ and $\Theta(x)$, then $G(x;t) = -\tilde{H}(t;x)$.*

For let ξ and η be any two distinct points of X and suppose $\xi < \eta$. Then we have the matrix equation

$$\begin{aligned} \int_a^b \{\tilde{H}(x;\xi)[(\partial/\partial x)G(x;\eta) - A(x)G(x;\eta)] \\ + [(\partial/\partial x)\tilde{H}(x;\xi) + \tilde{H}(x;\xi)A(x)]G(x;\eta)\}dx \\ = \tilde{H}(x;\xi)G(x;\eta)\Big|_{x=a}^{x=\xi-} + \tilde{H}(x;\xi)G(x;\eta)\Big|_{x=\xi+}^{x=\eta-} + \tilde{H}(x;\xi)G(x;\eta)\Big|_{x=\eta+}^{x=b} \end{aligned}$$

But

$$\int_a^b \tilde{H}(x;\xi)[(\partial/\partial x)G(x;\eta)-A(x)G(x;\eta)]dx$$

$$=-\int_a^b \tilde{H}(x;\xi)\Psi(x)R^r Z(\eta)\,dx,$$

$$=0,$$

and

$$\int_b^b [(\partial/\partial x)\tilde{H}(x;\xi)+\tilde{H}(x;\xi)A(x)]G(x;\eta)\,dx$$

$$=-\int_a^b {}^rY(\xi)\mathfrak{R}\Theta(x)G(x;\eta)\,dx,$$

$$=0.$$

Therefore

$$\tilde{H}(x;\xi)G(x;\eta)\,\Big|_{x=a}^{x=b}=\tilde{H}(x;\xi)G(x;\eta)\,\Big|_{x=\xi-}^{x=\xi+}+\tilde{H}(x;\xi)G(x;\eta)\,\Big|_{x=\eta-}^{x=\eta+}.$$

We have in view of (25), (36) and (42) that

$$[\tilde{H}(\xi+;\xi)-\tilde{H}(\xi-;\xi)]G(\xi;\eta)=-\tilde{H}(\eta;\xi)[G(\eta+;\eta)-G(\eta-;\eta)]$$

and hence in view of (35) and (41) that

$$G(\xi;\eta)=-\tilde{H}(\eta;\xi).$$

Since this is true for each pair of distinct points ξ and η, we have established Theorem 5.

COROLLARY. *For any chosen pair of matrices* $\Psi(x)$ *and* $\Theta(x)$ *satisfying the conditions described above, the principal generalized Green's matrix for the compatible system* (2), (3), *and also for the compatible adjoint system* (4), (5), *is unique.*

The matrices $\Psi(x)$ and $\Theta(x)$ have been chosen so that the r-rowed square matrices (28) and (39) have non-vanishing determinants. In particular, $\Psi(x)$ and $\Theta(x)$ may be determined so that each of the matrices (28) and (39) is the identity matrix. If this is done some of the preceding formulas will simplify considerably.

4. *Examples.* Consider the system

$$y'=\begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}\cdot y, \tag{44}$$

$$\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}\cdot y(-\pi)+\begin{Vmatrix} -1 & 0 \\ 0 & 1 \end{Vmatrix}\cdot y(\pi)=0. \tag{45}$$

This system is compatible of order one, and has a solution $y(x)=(\cos x, -\sin x)$. The adjoint system is

$$z' = -z \cdot \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}, \tag{46}$$

$$z(-\pi) \cdot \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} + z(\pi) \cdot \begin{Vmatrix} -1 & 0 \\ 0 & 1 \end{Vmatrix} = 0, \tag{47}$$

which has the solution $z(x) = (\cos x, -\sin x)$. According to the notation introduced above we have the matrices ${}^1Y(x)$ and ${}^1Z(x)$ defined by

$${}^1Y(x) = \begin{Vmatrix} \cos x & 0 \\ -\sin x & 0 \end{Vmatrix},$$

and

$${}^1Z(x) = \begin{Vmatrix} \cos x & -\sin x \\ 0 & 0 \end{Vmatrix}.$$

A matrix solution $Y(x)$ of (44) and its reciprocal $Z(x)$, which is a matrix solution of (46), is given by

$$Y(x) = \begin{Vmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{Vmatrix}, \qquad Z(x) = \begin{Vmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{Vmatrix}.$$

The generalized Green's matrix for the system (44), (45) which is given by the formula (16) is readily found to be equal to the matrix

$$K(x;t) = (1/2)Y(x)Z(t)[\,|x-t|/(x-t)].$$

From Theorem 2 it then follows that every generalized Green's matrix for (44), (45) is of the form

$$G(x;t) = K(x;t) + {}^1Y(x)U(t) + V(x){}^1Z(t), \tag{48}$$

where the matrices $U(t)$ and $V(x)$ depend upon the choice of the matrices $\Psi(x)$ and $\Theta(x)$.

(a) If we choose $\Psi(x) = (1/2\pi){}^1Y(x)$ and $\Theta(x) = (1/2\pi){}^1Z(x)$, then the principal generalized Green's matrix for the system (44), (45) is given by the relation (48), where

$$U(t) = 1/2\pi \begin{Vmatrix} t\cos t & -t\sin t \\ 0 & 0 \end{Vmatrix}, \qquad V(x) = 1/2\pi \begin{Vmatrix} -x\cos x & 0 \\ x\sin x & 0 \end{Vmatrix} = -\tilde{U}(x).$$

(b) If we choose

$$\Psi(x) = \Theta(x) = 1/\pi \begin{Vmatrix} \cos x & 0 \\ 0 & 0 \end{Vmatrix},$$

then the values of $U(t)$ and $V(x)$ become

$$U(t) = 1/2\pi \left\| \begin{matrix} t\cos t + \sin t & -t\sin t \\ 0 & 0 \end{matrix} \right\| = -\tilde{V}(t).$$

(c) If

$$\Psi(x) = 1/\pi \left\| \begin{matrix} \cos x & 0 \\ 0 & 0 \end{matrix} \right\|, \qquad \Theta(x) = 1/\pi \left\| \begin{matrix} 0 & -\sin x \\ 0 & 0 \end{matrix} \right\|,$$

the corresponding values of $U(t)$ and $V(x)$ are given by

$$U(t) = 1/2\pi \left\| \begin{matrix} t\cos t - \sin t & -t\sin t \\ 0 & 0 \end{matrix} \right\|,$$

$$V(x) = 1/2\pi \left\| \begin{matrix} -x\cos x - \sin x & 0 \\ x\sin x & 0 \end{matrix} \right\|.$$

For each choice of the matrices $\Psi(x)$ and $\Theta(x)$ the corresponding value of $H(x;t)$, the principal generalized Green's matrix for the adjoint system (46), (47), is determined by the relation $\tilde{H}(x;t) = -G(t;x)$, in view of Theorem 5.

5. *Remark.* The author has recently shown that for certain types of infinite systems of ordinary linear differential equations of the first order with two-point boundary conditions an ordinary Green's matrix may be defined, whenever the system is incompatible, in a manner entirely analogous to that used in the finite case.* For such infinite linear systems we may establish by the above method the existence of the principal generalized Green's matrix.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

* W. T. Reid, *loc. cit.* The infinite linear system considered may be written in the vector form (2), (3), where M and N are constant limited matrices of a certain type, $A(x)$ is an infinite matrix which is limited by a summable function on the interval X, and the solution $y(x)$ is a vector in Hilbert space. A case in which the solution $y(x)$ is a vector in a more general space than Hilbert space has also been considered in a second paper. See *Annals of Mathematics*, Vol. 32 (1931), p. 37.

A CREMONA GROUP OF ORDER THIRTY-TWO OF CUBIC TRANSFORMATIONS IN THREE-DIMENSIONAL SPACE.

By Ethel Isabel Moody.

1. *Introduction.* Since each four-dimensional quadric variety of the ∞^5 system $|H(z)| = 0$ having a common self-polar simplex Σ is transformed into itself by the G_{32} of linear transformations consisting of the six central harmonic homologies, defined by the vertices and opposite four-dimensional faces of Σ, and the different products of these, the intersections of these varieties by twos, threes, $\cdots$ are also invariant under G_{32}. By successive stereographic projections these varieties can be mapped in S_3. It is the purpose of this paper to derive the equations of the corresponding transformations in S_4 and in S_3 and to determine their characteristics and fundamental elements.

In a paper, published in the *Atti del R. Istituto Veneto,** Montesano has given a brief synthetic outline of this group of transformations, which he calls the group H, and also of the corresponding groups associated with two and three dimensional quadric varieties.

The corresponding G_4 in S_2 is expressed by the non-cyclic G_4 of harmonic homologies.

In S_3 the G_8 of central and axial involutions can be projected into a plane, where it is represented by four perspective quadratic inversions, and the three products of these in pairs. A cubic curve is left invariant; on it the three non-perspective inversions fix the three fundamental irrational involutions belonging to the curve.†

The G_{16} in S_4 can be mapped in S_3 by means of a group of sixteen quadratic transformations which transform into itself each surface of an ∞^3 system of $F_4 : C_2{}^2$. Within this system there are ∞^2 which are composite consisting of the plane of the conic and a cubic surface.‡ By mapping an F_3

* D. Montesano, *Su alcuni gruppi chiusi di trasformazioni involutorie nel piano e nello spazio*, Ser. 6, Vol. 6 (1888), pp. 1425-1444.

† The equations and properties of this group (Montesano's group γ) were derived by Miss B. I. Hart in her thesis for the degree of Master of Arts, Cornell University, September, 1926.

‡ The equations and essential properties of this group in S_3 (Montesano's group G) of quadratic transformations were derived in my thesis for the degree of Master of Arts, Cornell University, June, 1927.

of this system upon a plane, G_{16} becomes a group of plane cubic transformations, consisting of five perspective Jonquieres involutions and their products by twos. The group is completely determined when the positions of the five vertices of the perspective Jonquieres transformations are known, since the cubic curve passing through the vertex O_i of the transformation J_i and touching the lines joining O_i to the other four vertices at these vertices is uniquely determined. This cubic curve is point by point invariant under the transformation J_i.

2. *The G_{32} of linear transformations in S_5.* The equation of any four-dimensional quadratic variety in S_5 referred to a self-polar simplex as simplex of reference can be written in the form

$$H(z) \equiv \sum_{i=1}^{6} h_i z_i^2 = 0.$$

This variety is invariant under the group of thirty-two linear transformations, consisting of the six central harmonic homologies

$$T_i: \quad z_i' = -z_i, \quad z_r' = z_r; \qquad (r = 1 \cdots 6,\ r \neq i),$$
$$i = 1 \cdots 6$$

the fifteen products of these in pairs

$$T_{ij}: \quad z_i' = -z_i, \quad z_j' = -z_j, \quad z_r' = z_r; \quad (r = 1 \cdots 6,\ r \neq i, j),$$
$$i = 1 \cdots 5,\ j = 2 \cdots 6,\ i \neq j,$$

the ten products of these taken three at a time

$$T_{1ij}: \quad z_1' = -z_1, \quad z_i' = -z_i, \quad z_j' = -z_j, \quad z_r' = z_r; \quad (r = 2 \cdots 6,\ r \neq i, j),$$
$$i = 2 \cdots 5,\ j = 3 \cdots 6,\ i \neq j,\ i, j \neq 1$$

and the identity. The products of the T_i taken one, two, and three at a time will be referred to as transformations of the first, second, and third species respectively.

A transformation of the first species T_i is defined by the center

$$O_i: (z_i = 1,\ z_r = 0),\ (r = 1 \cdots 6,\ r \neq i)$$

and the four-space S_4 of invariant points

$$S_4: z_i = 0.$$

The defining elements of a transformation of the second species are the line O_iO_i and the three-space S_3 of invariant points

$$O_iO_j: \quad z_{r_1} = z_{r_2} = z_{r_3} = z_{r_4} = 0; \qquad S_3: \quad z_i = z_j = 0,$$

while those of a transformation of the third species are the two planes of invariant points

$$(O_1O_iO_j): \quad z_{r_1}=0, \quad z_{r_2}=0, \quad z_{r_3}=0,$$

and

$$S_2: \quad z_1=0, \quad z_i=0, \quad z_j=0.$$

The sections of $H(z)=0$ by the invariant S_4, S_3, and S_2 are the point by point invariant varieties Ω_i, Ω_{ij}, Ω_{1ij} of the transformations of the first, second, and third species respectively. These are all quadric varieties, Ω_i being three-dimensional, Ω_{ij} two-dimensional and Ω_{1ij} one-dimensional.

3. *The G'_{32} of quadratic transformations in S_4.* Let $(\bar{z})$ be a point of $H(z)=0$ and from it project $H(z)=0$ into $S_4: z_6=0$. The equations of this stereographic projection are

$$(3.1) \qquad y_g=\bar{z}_6 z_g - z_6 z_g; \qquad (g=1\cdots 5),$$

where the y's represent the coördinates of points in $z_6=0$. Conversely, given a point (y) in $z_6=0$, the coördinates of the point corresponding to it on $H(z)=0$ are

$$(3.2) \qquad \begin{aligned} z_g &= 2y_g H'(y,\bar{z}) - \bar{z}_g H'(y); \qquad (g=1\cdots 5) \\ z_6 &= -\bar{z}_6 H'(y) \end{aligned}$$

where

$$H'(y)=\sum_{g=1}^{5} h_g y_g^2$$

and

$$H'(y,\bar{z})=\sum_{g=1}^{5} h_g y_g \bar{z}_g.$$

The projection T' of T in $z_6=0$ is the product $(3.2)T(3.1)$, and the equations of the transformations of G'_{32} are

$$T_i': \quad (i=1\cdots 5) \qquad \begin{aligned} y_i' &= -y_i H'(y,\bar{z}) + \bar{z}_i H'(y) \\ y_r' &= y_r H'(y,\bar{z}); \qquad (r=1\cdots 5,\ r\neq i) \end{aligned}$$

$$T_6': \qquad y_r' = y_r H'(y,\bar{z}) - \bar{z}_r H'(y); \qquad (r=1\cdots 5)$$

$$T'_{ij}: \quad (i=1\cdots 5,\ j=2\cdots 5,\ i\neq j) \qquad \begin{aligned} y_i' &= -y_i H'(y,\bar{z}) + \bar{z}_i H'(y) \\ y_j' &= -y_j H'(y,\bar{z}) + \bar{z}_j H'(y) \\ y_r' &= y_r H'(y,\bar{z}); \qquad (r=1\cdots 5,\ r\neq i,j) \end{aligned}$$

$$T'_{i6}: \quad (i=1\cdots 5) \qquad \begin{aligned} y_i' &= -y_i H'(y,\bar{z}) \\ y_r' &= y_r H'(y,\bar{z}) - \bar{z}_r H'(y); \qquad (r=1\cdots 5,\ r\neq i) \end{aligned}$$

$$T'_{1ij}: \quad (i=2\cdots 4,\ j=3\cdots 5,\ i\neq j) \qquad \begin{aligned} y_1' &= -y_1H'(y,\bar z)+\bar z_1H'(y)\\ y_i' &= -y_iH'(y,\bar z)+\bar z_iH'(y)\\ y_j' &= -y_jH'(y,\bar z)+\bar z_jH'(y)\\ y_r' &= y_rH'(y,\bar z); \qquad (r=2\cdots 5,\ r\neq i,j)\end{aligned}$$

and

$$T'_{1i6}: \quad (i=2\cdots 5) \qquad \begin{aligned} y_1' &= -y_1H'(y,\bar z)\\ y_i' &= -y_iH'(y,\bar z)\\ y_r' &= y_rH'(y,\bar z)-\bar z_rH'(y); \qquad (r=2\cdots 5,\ r\neq i).\end{aligned}$$

The quadric surface

$$Q: \quad H'(y,\bar z)=0, \quad H'(y)=0$$

is a common fundamental element of all the transformations of G'_{32}.

The fundamental points of the transformations of G'_{32} are

$$\begin{aligned}
O_i'&: && (y_i=1, \quad y_r=0)\\
O_6'&: && (y_r=\bar z_r)\\
O'_{ij}&: && (y_i=\bar z_i, \quad y_j=\bar z_j, \quad y_r=0)\\
O'_{i6}&: && (y_i=0, \quad y_r=\bar z_r)\\
O'_{1ij}&: && (y_1=\bar z_1, \quad y_i=\bar z_i, \quad y_j=\bar z_j, \quad y_r=0)\\
O'_{1i6}&: && (y_1=y_i=0, \quad y_r=\bar z_r)
\end{aligned}$$

For each of these O' the restrictions on the i, j, and r are the same as for the T with corresponding subscript.

The image of the fundamental point O' under the corresponding transformation T' is the $S_3: H'(y,\bar z)=0$.

The image of Q under a transformation of the first species of G'_{32} is its three-dimensional projecting cone from O'. To prove this, let (p) be a point of Q and let the coördinates of O' be represented by Y. The coördinates of any point (y) on the line joining (p) to O' are of the form

$$y_g=\lambda Y_g+\mu p_g; \qquad (g=1\cdots 5).$$

By substituting these values for the y's in the equations of the transformations of the first species of G'_{32} and making use of the fact that $H'(p,\bar z)=0$ and $H'(p)=0$, the resulting point (y') is found to be (p). Hence each point of Q is imaged by a generator of the projecting cone of Q having its vertex at O'. This cone is of the second order. Its equation for T_i' is

$$[H'(y,\bar z)_i]^2+h_i\bar z_i^2H'(y)_i=0;*$$

and for T_6'

* A subscript written after the parenthesis will be used throughout this paper to indicate that the function to which it applies contains no term in the variable of that subscript.

$$H'(\bar{z})H'(y) - [H'(y,\bar{z})]^2 = 0.$$

This cone with $[H'(y,\bar{z})]^3$ forms the complete Jacobian of the defining web.

The Jacobian of the defining web of a transformation of the second or third species of G'_{32} consists of $[H'(y,\bar{z})]^3$ and a three-dimensional quadric variety containing the quadric surface Q, of which it is the image, simply. The equation of this quadric for T'_{ij} is

$$[H'(y,\bar{z})]^2 - 2(h_i y_i \bar{z}_i + h_j \bar{z}_j y_j)H'(y,\bar{z}) \\ + (h_i \bar{z}_i^2 + h_j \bar{z}_j^2)H'(y) = 0;$$

that for T'_{i6},

$$H'(y,\bar{z})[H'(y,\bar{z}) - 2H'(y,\bar{z})_i] + H'(y)H'(\bar{z})_i = 0;$$

that for T'_{1ij},

$$H'(y,\bar{z})[H'(y,\bar{z}) - 2h_1 y_1 \bar{z}_1 - 2h_i y_i \bar{z}_i - 2h_j y_j \bar{z}_j] \\ + H'(y)[h_1 \bar{z}_1^2 + h_i \bar{z}_i^2 + h_j \bar{z}_j^2] = 0;$$

and that for T'_{1i6},

$$H'(y,\bar{z})[H'(y,\bar{z}) - 2h_{r_1} y_{r_1} \bar{z}_{r_1} - 2h_{r_2} y_{r_2} \bar{z}_{r_2} - 2h_{r_3} y_{r_3} \bar{z}_{r_3}] \\ + H'(y)[h_{r_1} \bar{z}_{r_1} + h_{r_2} \bar{z}_{r_2} + h_{r_3} \bar{z}_{r_3}] = 0.$$

The projection upon $z_6 = 0$ of the point by point invariant quadric variety Ω of a transformation T of the first species of G_{32} is the invariant variety Ω' of the corresponding transformation T' of the first species of G'_{32}. The equations of the Ω''s are therefore

$$\Omega_i': \quad 2y_i H'(y,\bar{z}) - \bar{z}_i H'(y) = 0; \qquad (i = 1 \cdots 5)$$

and

$$\Omega_6': \quad H'(y) = 0.$$

The result of eliminating y_5 between Ω_i' and Ω_j' shows that the complete intersection of these two varieties lies in the three-space $y_i \bar{z}_j - y_j \bar{z}_i = 0$, and the intersection of this three-space with each of the quadrics Ω_i' and Ω_j' or

$$\begin{array}{ll} \Omega'_{ij} & 2\bar{z}_j y_j H'(y,\bar{z})_i - \bar{z}_j^2 H'(y)_i + h_i y_j^2 \bar{z}_i^2 = 0 \\ (i = 1 \cdots 5 & y_i \bar{z}_j - y_j \bar{z}_i = 0 \\ j = 2 \cdots 6 \quad i \neq j) & \end{array}$$

is the point by point invariant quadric surface of T'_{ij}; and

$$\Omega'_{i6}: \quad H'(y)_i = 0, \quad y_i = 0; \qquad (i = 1 \cdots 5)$$

is the point by point invariant surface of T'_{i6}. In a similar manner, the point by point invariant conic of T'_{1ij} is

$$\begin{array}{ll} \Omega'_{1ij}: & 2\bar{z}_j y_j H'(y,\bar{z})_{i,1} - \bar{z}_j^2 H'(y)_{i,1} + y_j^2(h_1 \bar{z}_1^2 + h_i \bar{z}_i^2) = 0 \\ (i = 2 \cdots 5, & y_1 \bar{z}_i \bar{z}_j = y_i \bar{z}_1 \bar{z}_j = y_j \bar{z}_1 \bar{z}_i. \\ j = 3 \cdots 6, \quad i \neq j, \quad i, j \neq 1) & \end{array}$$

The corresponding conic for T'_{1i6} is

$$\Omega'_{1i6}: \qquad H'(y)=0$$
$$(i=2\cdots 5) \quad y_1=y_i=0.$$

Given a point common to $H(z)=0$ and another quadric variety $G(z)=0$ of the system $|H(z)|=0$, the coördinates of its projection upon $z_6=0$ are

$$\rho y_i=\bar{z}_6 z_i-z_6\bar{z}_i; \qquad (i=1\cdots 5). \tag{3.4}$$

From (3.4) it follows that

$$z_i/z_6=(ky_i+\bar{z}_i)/\bar{z}_6. \tag{3.5}$$

If the values of the ratios in (3.5) are substituted in $H(z)=0$ and $G(z)=0$, the resulting equations are

$$\begin{aligned} G(\bar{z})+2kG'(y,\bar{z})+k^2G'(y)&=0\\ 2H'(y,\bar{z})+kH'(y)&=0. \end{aligned} \tag{3.6}$$

The result of eliminating k between the equations (3.6) is

$$G(\bar{z})[H'(y)]^2-4G'(y,\bar{z})H'(y)H'(y,\bar{z})+4G'(y)[H'(y,\bar{z})]^2=0.$$

This equation represents a three-dimensional quartic variety in S_4 having the quadric surface

$$F_2:\ H'(y)=0, \quad H'(y,\bar{z})=0$$

as double quadric. The three-space $H'(y,z)=0$ of the double quadric is the polar three-space of O_i' with respect to Ω_i'.

Every three-dimensional quartic variety having a double quadric is invariant under a G'_{32}, for such a variety is the projection in S_4 of the intersection of two quadric varieties in S_5 which have a common self-polar simplex.

Let $A(z)=0$ be a second quadric variety belonging to the system $|H(z)|=0$ and containing the center of projection $(\bar{z})$. The coördinates of any point (z') on the line joining a point (z) to $(\bar{z})$ are

$$z_i'=\lambda\bar{z}_i+\mu z_i; \qquad (i=1\cdots 6).$$

The points of intersection of this line with $A(z)=0$ and $H(z)=0$ are given by

$$2\lambda H(z,\bar{z})+\mu H(z)=0, \qquad 2\lambda A(z,\bar{z})+\mu A(z)=0.$$

The condition that these two equations in λ and μ have a solution in common is

$$H(z,\bar{z})A(z)-A(z,\bar{z})H(z)=0.$$

This is the equation of the projecting cone with vertex at $(\bar{z})$ of the intersection of these two varieties. The equation of the projection of this intersection upon $S_4: z_6=0$ is therefore

$$V_3: \quad H'(y,\bar{z})A'(y) - A'(y,\bar{z})H'(y) = 0,$$

which is the equation of a three-dimensional cubic variety containing a plane. The equations of this plane are

$$H'(y,\bar{z}) = 0, \qquad A'(y,\bar{z}) = 0.$$

But, in general, the projection upon S_4 of the intersection of two quadric varieties is a quartic variety. The residual part of this projection is the three-space corresponding to $(\bar{z})$ in the stereographic projection. This three-space is

$$H'(y,\bar{z}) = 0,$$

the intersection of the tangent S_4 to

$$H(z) = 0$$

at $(\bar{z})$ and the S_4 of projection. This is the three-space of the double quadric of V_4.

Therefore, among the ∞^4 quartic varieties which are invariant under G'_{32} there are ∞^3 which are composite, consisting of the cubic variety and the three-space of the double quadric.

4. *G''_{32} of cubic transformations in S_3.* The plane

$$A'(y,\bar{z}) = 0, \qquad H'(y,\bar{z}) = 0$$

meets the quartic surface

$$A'(y) = 0, \qquad H'(y) = 0$$

in four points which are double points of

$$V_3: \quad H'(y,\bar{z})A'(y) - A'(y,\bar{z})H'(y) = 0.^*$$

V_3 can be projected stereographically into $S_3: y_5 = 0$ from any one $(\bar{y})$ of these four double points. The equations of the projection are

$$x_g = y_g\bar{y}_5 - y_5\bar{y}_g; \qquad (g = 1 \cdots 4). \tag{4.1}$$

Conversely, given a point (x) in $y_5 = 0$, the point corresponding to it on V_3 is

$$\begin{aligned} y_g &= w\bar{y}_g + kx_g; \qquad (g = 1 \cdots 4) \\ y_5 &= w\bar{y}_5, \end{aligned} \tag{4.2}$$

where

$$\begin{aligned} k &= -2\,[H''(x,\bar{z})A''(x,\bar{y}) - A''(x,\bar{z})H''(x,\bar{y})] \\ w &= H''(x,\bar{z})A''(x) - A''(x,\bar{z})H''(x).\dagger \end{aligned}$$

---

* The properties of a three-dimensional cubic variety containing a plane are discussed by Segre, "Sulle varietà cubiche dello spazio a quattro dimensioni e su certi sistemi di rette e certe superficie dello spazio ordinario." See sections 5-11. *Memorie della Reale Accademia delle Scienze di Torino*, Ser. 2, Vol. 39 (1889), pp. 3-48.

† The H'' and A'' are used to denote functions similar to the H''s of equations (3.2) but for which the summations are from $g = 1$ to $g = 4$.

The projection T'' of a transformation T' in $y_5 = 0$ is the product $(4.2)(T')(4.1)$. Let

$$m = A''(x)H''(x, \bar{y}) - A''(x, \bar{y})H''(x).$$

The equations of the transformations T'' can then be written in the form

$$\begin{array}{lll}
T_i'': & x_i' = -kx_i - 2w\bar{y}_i + 2\bar{z}_i m & \\
(i = 1 \cdots 4) & x_r' = kx_r; & (r = 1 \cdots 4,\quad r \neq i). \\
T_5'': & x_r' = kx_r\bar{y}_5 + 2w\bar{y}_5\bar{y}_r - 2\bar{z}_5\bar{y}_r m; & (r = 1 \cdots 4). \\
T_6'': & x_r' = x_r\bar{y}_5 k - 2(\bar{z}_r\bar{y}_5 - \bar{z}_5\bar{y}_r)m; & (r = 1 \cdots 4). \\
T''_{ij}: & x_i' = -kx_i - 2w\bar{y}_i + 2\bar{z}_i m & \\
(i = 1 \cdots 3 & x_j' = -kx_j - 2w\bar{y}_j + 2\bar{z}_j m & \\
j = 2 \cdots 4 & x_r' = kx_r; & (r = 1 \cdots 4,\quad r \neq i, j). \\
i \neq j) & & \\
T''_{i5}: & x_i' = -kx_i\bar{y}_5 + 2(\bar{z}_i\bar{y}_5 - \bar{z}_5\bar{y}_i)m & \\
(i = 1 \cdots 4) & x_r' = kx_r\bar{y}_5 + 2\bar{y}_r\bar{y}_5 w - 2\bar{z}_5\bar{y}_r m; & (r = 1 \cdots 4,\quad r \neq i). \\
T''_{i6}: & x_i' = -kx_i\bar{y}_5 - 2w\bar{y}_i\bar{y}_5 + 2\bar{z}_5\bar{y}_i m & \\
(i = 1 \cdots 4) & x_r' = kx_r\bar{y}_5 - 2(\bar{z}_r\bar{y}_5 - \bar{z}_5\bar{y}_r)m; & (r = 1 \cdots 4,\quad r \neq i). \\
T''_{56}: & x_r' = kx_r + 2w\bar{y}_r - 2\bar{z}_r m; & (r = 1 \cdots 4). \\
T''_{1ij}: & x_1' = -kx_1 - 2w\bar{y}_1 + 2\bar{z}_1 m & \\
(i = 2 \cdots 4 & x_i' = -kx_i - 2w\bar{y}_i + 2\bar{z}_i m & \\
j = 3 \cdots 4 & x_j' = -kx_j - 2w\bar{y}_j + 2\bar{z}_j m & \\
i \neq j & x_r' = kx_r; & (r = 2 \cdots 4,\quad r \neq i, j). \\
i, j \neq 1) & & \\
T''_{1i5}: & x_1' = -kx_1\bar{y}_5 + 2(\bar{z}_1\bar{y}_5 - \bar{z}_5\bar{y}_1)m & \\
(i = 2 \cdots 4) & x_i' = -kx_i\bar{y}_5 + 2(\bar{z}_i\bar{y}_5 - \bar{z}_5\bar{y}_i)m & \\
 & x_r' = kx_r\bar{y}_5 + 2w\bar{y}_r\bar{y}_5 - 2\bar{z}_5\bar{y}_r m; & (r = 2 \cdots 4,\quad r \neq i). \\
T''_{1i6}: & x_1' = -kx_1\bar{y}_5 - 2w\bar{y}_5\bar{y}_1 + 2\bar{y}_1\bar{z}_5 m & \\
(i = 2 \cdots 4) & x_i' = -kx_i y_5 - 2w\bar{y}_5\bar{y}_i + 2\bar{y}_i\bar{z}_5 m & \\
 & x_r' = k\bar{y}_5 x_r - 2m(\bar{z}_r\bar{y}_5 - \bar{z}_5\bar{y}_r); & (r = 2 \cdots 4,\quad r \neq i). \\
T''_{156}: & x_1' = -kx_1 & \\
 & x_r' = kx_r + 2w\bar{y}_r - 2\bar{z}_r m; & (r = 2 \cdots 4).
\end{array}$$

The right hand members of these transformations are of the third degree and will all vanish when $-k/2 = 0$, $w = 0$, and $m = 0$. But each of these equations is exactly the condition that the other two, considered as equations in

$$A''(x),\ H''(x);\quad H''(x, \bar{y}),\ A''(x, \bar{y});\quad \text{and}\quad H''(x, \bar{z}),\ A''(x, \bar{z})$$

respectively, have a solution in common. Therefore, if $k=0$ and $w=0$ when $H''(x,\bar{z})$ and $A''(x,\bar{z})$ are not both zero, $m=0$. But $k=0$ and $w=0$ intersect in a C_6, which is composite, consisting of the straight line

$$C_1:\quad H''(x,\bar{z})=0,\quad A''(x,\bar{z})=0$$

and a C_5. This C_5 is therefore a fundamental curve of all of the transformations of G''_{32}.

The fundamental points O'' of the transformations T'' of the first species are:

$$\begin{aligned} O_i'':&\quad (x_i=1,\quad x_r=0)\\ O_5'':&\quad (x_r=\bar{y}_r)\\ O_6'':&\quad (x_r=\bar{z}_r\bar{y}_5-\bar{z}_5\bar{y}_r). \end{aligned}$$

The point O'' is a double point of each of the surfaces of the defining web of the corresponding T'', and hence the six transformations of the first species of G''_{32} are monoidal. The fundamental points O'' all lie on the common fundamental C_5.

If (p) represents a point on this C_5 and if the point O'' is represented by (x'), the coördinates of any point (x) on a line joining (p) to (x') are

$$(4.3)\qquad x_g=\lambda p_g+\mu x_g'; \qquad (g=1\cdots 4).$$

If the x's in the equations of the transformations of the first species of G''_{32} are replaced by their values from (4.3), the image of any point on this line is found to be the point (p). Therefore, the image of C_5 under a transformation of the first species of G''_{32} is the quartic cone having the corresponding O'' as vertex and containing C_5.

The image of O'' under its corresponding T'' is the quadric $k=0$. Both C_1 and C_5 lie on $k=0$. The two systems of generators of this quadric are

$$(1)\qquad \begin{aligned} A''(x,\bar{z})-\mu H''(x,\bar{z})=0\\ A''(x,\bar{y})-\mu H''(x,\bar{y})=0 \end{aligned}$$

and

$$(2)\qquad \begin{aligned} A''(x,\bar{z})-\lambda A''(x,\bar{y})=0\\ H''(x,\bar{z})-\lambda H''(x,\bar{y})=0. \end{aligned}$$

C_1 is a line of (2), and, therefore, its symbol on $k=0$ is [1, 0]. The symbol of the complete intersection of $k=0$ and $w=0$ is [3, 3], and therefore, C_5 must be of symbol [2, 3] on $k=0$. Hence, through each point O'', passes one line of the regulus (2) and therefore one trisecant of C_5. This line is a fundamental element of the transformation to which O'' belongs, each point

of the line having for its image the whole line. The equations of these fundamental lines are:

L_i:

$$\bar{y}_i H''(x,\bar{z}) - \bar{z}_i H''(x,\bar{y}) = 0$$
$$\bar{y}_i A''(x,\bar{z}) - \bar{z}_i A''(x,\bar{y}) = 0$$

L_5:

$$H''(x,\bar{z})H''(\bar{y}) - H''(x,\bar{y})H''(\bar{y},\bar{z}) = 0$$
$$A''(x,\bar{z})H''(\bar{y}) - A''(x,\bar{y})H''(\bar{y},\bar{z}) = 0$$

L_6:

$$H''(x,\bar{z})[\bar{y}_5 H''(\bar{y},\bar{z}) - \bar{z}_5 H''(\bar{y})] - H''(x,\bar{y})[\bar{y}_5 H''(\bar{z}) - \bar{z}_5 H''(\bar{y},\bar{z})] = 0$$
$$A''(x,\bar{z})[\bar{y}_5 H''(\bar{y},\bar{z}) - \bar{z}_5 H''(\bar{y})] - A''(x,\bar{y})[\bar{y}_5 H''(\bar{z}) - \bar{z}_5 H''(\bar{y},\bar{z})] = 0.$$

The complete Jacobian of a transformation of the first species consists of k^2 and the quartic cone.

The invariant surfaces of the transformations of the first species of G''_{32} are found by projecting the Ω''s into $y_5 = 0$. Their equations are:

$$\Omega_i'': \quad w\bar{y}_i + kx_i - \bar{z}_i m = 0$$
$$\Omega_5'': \quad w\bar{y}_5 - \bar{z}_5 m = 0$$
$$\Omega_6'': \quad m = 0.$$

The general plane of the system through O_i'' is

$$\lambda x_{r_1} + \mu x_{r_2} + \nu x_{r_3} = 0.$$

Under T_i'' this becomes

$$k(x_{r_1}\lambda + \mu x_{r_2} + \nu x_{r_3}) = 0.$$

But $k = 0$ is the image of O'', and, therefore, the plane is transformed into itself. Any line l in such a plane, π, goes into a cubic curve under T_i''. Since the line l meets k in two points this C_3 must have a double point at O_i''. But, if a line l meets the fundamental trisecant L_i, the trisecant itself is a component of the image, and the line goes into a conic. But l has one other point S in common with $k = 0$ and meets each of the lines joining the point O'' to the two residual intersections R_1, R_2, of C_5 and π. Therefore, the image conic passes through O_i'', R_1 and R_2. A similar argument holds for T_5'' and T_6''. Therefore, in any plane of the pencil through L, the corresponding T'' of the first species of G''_{32} is quadratic, the image of any straight line l being a conic through R_1, R_2, and O''.

For transformations of the second species, the fundamental elements which are transformed into $k = 0$ are the straight lines:

$$
\begin{array}{lll}
O_i''O_j'': & x_{r_1}=0; & (r_1=1\cdots 4) \\
& x_{r_2}=0; & (r_2=1\cdots 4,\quad r_1,r_2\neq i,j,\quad r_1\neq r_2). \\
O_i''O_5'': & x_{r_1}\bar{y}_{r_3}-x_{r_3}\bar{y}_{r_1}=0 & \\
& x_{r_1}\bar{y}_{r_2}-x_{r_2}\bar{y}_{r_1}=0 & \\
& \quad (r_1=1\cdots 4,\quad r_2=1\cdots 4,\quad r_3=1\cdots 4) & \\
& \qquad (r_1\neq r_2\neq r_3;\quad r_1,r_2,r_3\neq i). & \\
O_i''O_6'': & x_{r_1}(\bar{z}_{r_2}\bar{y}_5-\bar{z}_5\bar{y}_{r_2})-x_{r_2}(\bar{z}_{r_1}\bar{y}_5-\bar{z}_5\bar{y}_{r_1})=0 & \\
& x_{r_1}(\bar{z}_{r_3}\bar{y}_5-\bar{z}_5\bar{y}_{r_3})-x_{r_3}(\bar{z}_{r_1}\bar{y}_5-\bar{z}_5\bar{y}_{r_1})=0 & \\
& \quad (r_1=1\cdots 4,\quad r_2=1\cdots 4,\quad r_3=1\cdots 4) & \\
& \qquad (r_1\neq r_2\neq r_3;\quad r_1,r_2,r_3\neq i). & \\
O_5''O_6'': & x_1(\bar{y}_2\bar{z}_3-\bar{z}_2\bar{y}_3)+x_2(\bar{y}_3\bar{z}_1-\bar{z}_3\bar{y}_1)+x_3(\bar{y}_1\bar{z}_2-\bar{z}_1\bar{y}_2)=0 & \\
& x_2(\bar{y}_3\bar{z}_4-\bar{z}_3\bar{y}_4)+(\bar{y}_4\bar{z}_2-\bar{z}_4\bar{y}_2)x_3+x_4(\bar{y}_2\bar{z}_3-\bar{z}_2\bar{y}_3)=0. &
\end{array}
$$

The cubic surface

$$\bar{z}_j(w\bar{y}_i+kx_i)-\bar{z}_i(w\bar{y}_j+kx_j)=0$$

meets each of the cubics Ω_i' and Ω_j' in the

$$C_9:\quad \bar{z}_j(w\bar{y}_i+kx_i)-\bar{z}_i(w\bar{y}_j+kx_j)=0,\qquad m=0.$$

This C_9 is the complete intersection of the two cubics Ω_i'' and Ω_j'', but it is composite, consisting of the C_5 and a residual C_4. This C_4 is therefore point by point invariant under the transformations T''_{ij}. The equations of these invariant C_4 of the transformations of the second species of G''_{32} are

$$
\begin{array}{ll}
C_{4_{ij}}: & (\bar{z}_j\bar{y}_i-\bar{z}_i\bar{y}_j)A''(x)-2(\bar{z}_jx_i-x_j\bar{z}_i)A''(x,\bar{y})=0 \\
& (\bar{z}_j\bar{y}_i-\bar{z}_i\bar{y}_j)H''(x)-2(\bar{z}_jx_i-x_j\bar{z}_i)H''(x,\bar{y})=0 \\
C_{4_{i5}}: & (\bar{y}_i\bar{z}_5-\bar{z}_i\bar{y}_5)A''(x)-2x_i\bar{z}_5A''(x,\bar{y})=0 \\
& (\bar{y}_i\bar{z}_5-\bar{z}_i\bar{y}_5)H''(x)-2x_i\bar{z}_5H''(x,\bar{y})=0 \\
C_{4_{i6}}: & \bar{y}_iA''(x)-2x_iA''(x,\bar{y})=0 \\
& \bar{y}_iH''(x)-2x_iH''(x,\bar{y})=0 \\
C_{4_{56}}: & A''(x)=0 \\
& H''(x)=0.
\end{array}
$$

If r is the rank of a space curve C_m, and C_m and $C_{m'}$ form the complete intersection of two surfaces, F_μ and F_ν, then

$$m(\mu+\nu-2)=r+t,$$

where t is the number of intersections of C_m and $C_{m'}$. C_4 and C_5 form the complete intersection of two cubics. The C_5 was of symbol $[2,3]$ on $k=0$, and the number of apparent double points of a curve of symbol $[k_1,k_2]$ is $\frac{1}{2}(k_1^2+k_2^2-k_1-k_2)$. Therefore C_5 has four apparent double points and is of genus 2. $r=m(m-1)-2h$, and therefore C_5, $p=2$, is of rank 12, and $t=8$. Therefore C_5 and C_4 have 8 points in common.

The fundamental lines of the transformations of the third species are:

$$L_{1ij}: \quad x_1(\bar{y}_i\bar{z}_j - \bar{z}_i\bar{y}_j) + x_i(\bar{y}_j\bar{z}_1 - \bar{y}_1\bar{z}_j) + x_j(\bar{y}_1\bar{z}_i - \bar{y}_i\bar{z}_1) = 0, \quad x_r = 0$$

$$L_{1i5}: \quad x_1(\bar{z}_i\bar{y}_5 - \bar{z}_5\bar{y}_i) - x_i(\bar{z}_1 y_5 - \bar{z}_5\bar{y}_1) = 0, \quad x_{r_1}\bar{y}_{r_2} - x_{r_2}\bar{y}_{r_1} = 0$$

$$L_{1i6}: \quad x_{r_1}(\bar{z}_{r_2}\bar{y}_5 - \bar{z}_5\bar{y}_{r_2}) - x_{r_2}(\bar{z}_{r_1}\bar{y}_5 - \bar{z}_5\bar{y}_{r_1}) = 0, \quad x_1\bar{y}_i - x_i\bar{y}_1 = 0$$

$$L_{156}: \quad x_1 = 0, \quad x_{r_1}\bar{y}_{r_2}\bar{z}_{r_3} + x_{r_3}\bar{y}_{r_1}\bar{z}_{r_2} + x_{r_2}\bar{y}_{r_3}\bar{z}_{r_1} - x_{r_3}\bar{y}_{r_2}\bar{z}_{r_1} - x_{r_1}\bar{y}_{r_3}\bar{z}_{r_2} - x_{r_2}\bar{y}_{r_1}\bar{z}_{r_3} = 0.$$

The invariant elements of the transformations of the third species are isolated points, for the

$$C_9: \quad \bar{z}_1(w\bar{y}_1 + kx_1) = \bar{z}_i(w\bar{y}_i + kx_i) = (w\bar{y}_j + kx_j)\bar{z}_j$$

meets each of the cubic surfaces Ω_1'', Ω_i'', Ω_j'' in the configuration of points represented by the equations

$$\bar{z}_1(w\bar{y}_1 + kx_1) = \bar{z}_i(w\bar{y}_i + kx_i) = \bar{z}_j(w\bar{y}_j + kx_j), \qquad m = 0.$$

These would in general represent 27 points, but the equivalence of C_5, $p = 2$, is 23, and, therefore, there are four residual points of intersection which are invariant under T''_{1ij}. The four points which with C_5 form the complete intersection of

$$\bar{z}_r(w\bar{y}_r + kx_r) = w\bar{z}_5\bar{y}_5 = 0, \qquad m = 0$$

are also invariant under T''_{1ij}. The invariant points of T''_{1i5} are those which with C_5 form the complete intersections of

$$\bar{z}_1(w\bar{y}_1 + kx_1) = \bar{z}_i(w\bar{y}_i + kx_i) = \bar{z}_5\bar{y}_5 w, \qquad m = 0$$

and of

$$\bar{z}_{r_1}(w\bar{y}_{r_1} + kx_{r_1}) = \bar{z}_{r_2}(w\bar{y}_{r_2} + kx_{r_2}) = 0, \qquad m = 0.$$

Those of T''_{1i6} are those which with C_5 form the complete intersection of

$$\bar{z}_1(w\bar{y}_1 + kx_1) = \bar{z}_i(w\bar{y}_i + kx_i) = 0, \qquad m = 0$$

and of

$$\bar{z}_{r_1}(w\bar{y}_{r_1} + kx_{r_1}) = \bar{z}_{r_2}(w\bar{y}_{r_2} + kx_{r_2}) = w\bar{z}_5\bar{y}_5, \qquad m = 0,$$

and those of T''_{156} are those which with C_5 form the complete intersection of

$$\bar{z}_1(w\bar{y}_1 + kx_1) = \bar{z}_5\bar{y}_5 w = 0, \qquad m = 0,$$

and of

$$\bar{z}_{r_1}(w\bar{y}_{r_1} + kx_{r_1}) = \bar{z}_{r_2}(w\bar{y}_{r_2} + kx_{r_2}) = \bar{z}_{r_3}(w\bar{y}_{r_3} + kx_{r_3}), \qquad m = 0.$$

The transformations of the first species of G''_{32} are monoidal, while those of the second and third species are not. The various types of non-

monoidal cubic involutorial transformations have been discussed by F. R. Morris.* In the work which follows, frequent reference has been made to his paper.

If a non-monoidal cubic transformation is to be involutorial, all of the cubics of the defining web must have a C_6, $p=3$ in common, and the image of a point P may be defined as the point of intersection of the polar planes of P with respect to all the quadrics of a bundle having C_6 as the locus of the vertices of the cones of the bundle. If this bundle of quadrics contains a composite quadric, the line of intersection of the component planes of this quadric must be a component of the vertex locus, and the whole locus consists of C_5, $p=2$, and one of its bisecants. This composite sextic is the fundamental curve of the corresponding non-monoidal cubic transformation.

The transformations of the second and third species of G''_{32} are of this type, and, from the properties of the general non-monoidal cubic transformations in S_3 having for their fundamental curves C_6's consisting of C_5, $p=2$, and one of its bisecants, the nature of the principal systems of these transformations can be determined. For the case in which C_6 is non-composite, its image consists of the ∞^1 trisecants of C_6, which constitute a ruled surface $R_8:C_6^3$. If, however, the C_6 is composite, consisting of a C_5 and a line, this R_8 breaks up into two parts, one consisting of the bisecants of C_5 which meet the fundamental line and the other of the trisecants of C_5. The trisecants of C_5 form one regulus of a quadric containing C_5. This quadric is the image of the fundamental line. The bisecants of C_5 which meet the fundamental line generate a ruled surface of order six, containing C_5 as a double curve and the fundamental bisecant as triple line. This sextic surface is the image of C_5. Therefore, for a transformation of either the second or third species, the principal system consists of the quadric surface, $k=0$, image of the fundamental line, and the image of C_5 which is a ruled sextic surface, containing the fundamental line as triple line and the C_5 as double curve.

Any plane through $O_i''O_j''$ has an equation of the form

$$\lambda x_{r_1}+\mu x_{r_2}=0.$$

Its image under T'''_{ij} is

$$k(\lambda x_{r_1}+\mu x_{r_2})=0.$$

The image of $O_i''O_j''$ is $k=0$, and therefore under T'''_{ij} every plane π of the pencil through $O_i''O_j''$ is transformed into itself. Under T_i'' the image of any straight line l of π is a cubic curve having a double point at O_i''. Since l meets each of the lines joining O_i'' to the three residual points of intersection R_1, R_2, R_3 of C_5 and π, the image cubic must pass through these

* Morris, "Classification of Involutory Cubic Space Transformations," *University of California Publications in Mathematics*, Vol. 1, No. 11 (1920), pp. 223-240.

points. Under T_j'' this cubic goes into a C_9. But, since C_3 contains O_j'', $k\pi$ (the section of $k=0$ by π) must be a component of this C_9, as must also the lines joining O_j to the three residual intersections of C_5 and π; and the line $O_i''O_j''$, counted twice, is also a component. Therefore, the image of l is a conic containing R_1, R_2, R_3. The same argument holds for planes through the lines $O_i''O_5''$, $O_i''O_6''$, and $O_5''O_6''$. Therefore, in any plane π of the pencil through a fundamental line, the corresponding transformation of the second species is quadratic, the image of any line l being a conic containing R_1, R_2, and R_3.

The equation of any plane through the line L_{1ij} is

$$\lambda x_1(\bar{y}_i\bar{z}_j-\bar{y}_j\bar{z}_i)+\lambda x_i(\bar{y}_j\bar{z}_1-\bar{z}_j\bar{y}_1)+\lambda x_j(\bar{y}_1\bar{z}_i-\bar{y}_i\bar{z}_1)+\mu x_r=0.$$

Under T''_{1ij} this goes into

$$k[\lambda x_1(\bar{y}_i\bar{z}_j-\bar{y}_j\bar{z}_i)+\lambda x_i(\bar{y}_j\bar{z}_1-\bar{y}_1\bar{z}_j)+\lambda x_j(\bar{y}_1\bar{z}_i-\bar{y}_i\bar{z}_1)-\mu x_r]=0.$$

The component $k=0$ is the image of L_{1ij}, and any plane through L_{1ij} is transformed into a plane through L_{1ij}. The particular plane of the system through L_{1ij} which contains O_1'', O_i'', O_j'' is $x_r=0$, and this is transformed into itself, as is also the plane containing O_r'', O_5'', O_6'' or

$$x_1(\bar{y}_i\bar{z}_j-\bar{y}_j\bar{z}_i)+x_i(\bar{y}_j\bar{z}_1-\bar{y}_1\bar{z}_j)+x_j(\bar{y}_1\bar{z}_i-\bar{y}_i\bar{z}_1)=0.$$

Let the two points common to L_{1ij} and C_5 be R_1 and R_2. Then in the plane determined by O_1'', O_i'', O_j'', any straight line l goes into $C_3: O_iO_jR_1R_2O_1^2$ under T_1''. Under T_i''

$$C_3:\ O_iO_jR_1R_2O_1^2\sim C_2:\ R_1R_2O_j.$$

Under T_j''

$$C_2:\ R_1R_2O_j\sim C_2:\ O_1O_iO_j.$$

Therefore, in this plane, the transformation T''_{1ij} is quadratic, any line being conjugate to a conic through O_1, O_i, and O_j. A similar argument holds for the plane through L_{1ij} and O_r'', O_5'', O_6'', and also for the corresponding planes through the other fundamental lines L and their corresponding transformations. Hence, associated with each transformation of the third species of G''_{32} are two planes which are transformed each into itself and in which the transformation is quadratic.

The projection in $y_5=0$ of the intersection of the given V_3 and another V_3 of the system

$$H'(y,\bar{z})B'(y)-B'(y,\bar{z})H'(y)=0$$

is

$$\begin{aligned} F_7:\quad & w^2[H''(x,\bar{z})B''(\bar{y})-2B''(y,\bar{z})H''(x,\bar{y})]\\ &+kw[2H''(x,\bar{z})B''(\bar{y},x)-B''(\bar{y},\bar{z})H''(x)-2B''(x,\bar{z})H''(x,\bar{y})]\\ &+k^2[H''(x,\bar{z})B''(x)-B''(x,\bar{z})H''(x)]=0\end{aligned}$$

on which $k=0, \quad w=0,$

is double. But $H''(x, \bar{z})=0$, the image of $(\bar{y})$, is a component of F_7. The other component is

$$F_6: \quad B''(\bar{y})w^2 - 2B''(\bar{y}, \bar{z})wm + 2kwB''(\bar{y}, x) \\ - 2B''(x, \bar{z})km + B''(x)k^2 = 0.$$

The double line

$$C_1: \quad H''(x, \bar{z})=0, \quad A''(x, \bar{z})=0$$

lies in the plane $H''(x, \bar{z})=0$, but the double C_5 does not, and hence there are ∞ $F_6: C_5{}^2, C_1$, which are invariant under G''_{32}.

Every cubic variety of the ∞^3 system

$$H'(y, \bar{z})B'(y) - B'(y, \bar{z})H'(y) = 0$$

is invariant under G'_{32}. The cubic variety

$$H'(y, \bar{z})A'(y) - A'(y, \bar{z})H'(y) = 0$$

has $(\bar{y})$ as double point, and every variety of the system contains $(\bar{y})$ simply. A variety of the system will have $(\bar{y})$ as a double point provided $B'(\bar{y})=0$ and $B'(\bar{y}, \bar{z})=0$. Therefore, there are ∞^1 cubics of the system which have $(\bar{y})$ as a double point. Let

$$H'(y, \bar{z})C'(y) - C'(y, \bar{z})H'(y) = 0$$

represent one such variety. The projection in $S_3: y_5=0$ of the intersection of this variety and

$$H'(y, \bar{z})A'(y) - A'(y, \bar{z})H'(y) = 0$$

is composite, consisting of

$$H''(x, \bar{z}) = 0$$

and the

$$F_4: \quad kC''(x) - 2mC''(x, \bar{z}) + 2wC''(x, \bar{y}) = 0.$$

Therefore, there are ∞^0 $F_4: C_5$ which are invariant under G''_{32}. The F_4 have all six of the points O'' as double points and contain the fifteen lines joining these in pairs. The quadric cone K_2, determined by a vertex O'' and the five lines of F_4 which pass through it, meets F_4 in a C_8 which is composite, consisting of the five lines and the C_3 determined by the six points O''. The F_4 must therefore contain this C_3. An F_4 having these properties is a Weddle surface, and, since the Weddle surface is known to be irrational, and F_4 and $[H''(x, \bar{z})]^2=0$ form a composite F_6 of the system of surfaces invariant under G''_{32}, the general sextic of this system cannot be mapped upon a plane.

CORNELL UNIVERSITY.

ON THE CAPACITY OF SETS OF CANTOR TYPE.*

By Oliver D. Kellogg.

1. *Introduction.* If E denotes any bounded set of points, then E, together with its limit points, contains the boundary of an infinite domain T. To this domain and to the boundary values 1, may be assigned, by the method of sequences,† a harmonic function, called the *conductor potential* of the set E. This conductor potential, V, is unique, in the sense that it is the only one yielded by the method of sequences, and in particular, is independent of the sequence of regions with T as limit. The points of E at which V approaches 1 (and also points of E not boundary points of T), if such exist, are called *regular* points of E. All other points of E are called *exceptional.* The total mass, as given by Gauss' integral, producing the potential V, is called the capacity of E.

The question of the unique determination of a harmonic function by continuous boundary values, not yet generally settled for boundaries containing exceptional points, would be definitely and affirmatively answered if the following lemma were established:

Any bounded closed set of points of positive capacity contains regular points.

This lemma was formulated in 1926,‡ and established in the case of the logarithmic potential in 1928,§ but it has not yet been proved in the case of the Newtonian potential. In view of this fact, and of the central position of the lemma with respect to the Dirichlet problem for general domains, the consideration of the lemma in special cases appears to be a task worth while.

In particular, the sets of Cantor type, studied in the following pages, have a peculiar interest, for the following reason. Some of them have 0 capacity, and some positive capacity. The latter owe their positive capacity to the relatively great separation of the points of the set, a quality which ordinarily reduces the likelihood of regular points. It would accordingly

* Read before the American Mathematical Society, September 11, 1930.

† See Kellogg and Vasilesco, "A Contribution to the Theory of Capacity," *American Journal of Mathematics*, Vol. 51 (1929), pp. 515-526. References to the literature are there given.

‡ Kellogg, "On the Classical Dirichlet Problem for General Domains," *Proceedings of the National Academy of Sciences*, Vol. 12 (1926), p. 406, foot-note 11.

§ Kellogg, "Unicité des fonctions harmoniques," *Comptes Rendus*, Vol. 13 (1928), pp. 526-27.

seem probable that if the lemma were to fail, it would fail for a set of this sort. In the general cases studied, it is true.

2. *Construction of the Sets of Cantor Type.* Let C denote a closed cube, of unit side, and let $\alpha_1, \alpha_2, \alpha_3, \cdots$ be an infinite sequence of positive proper fractions. We remove from C all points whose distances from any of the three planes through the center, and parallel to the faces, are less than $\alpha_1/2$. There remain eight closed cubes of side $\delta_1 = (1 - \alpha_1)/2$. From each of these, we remove all points whose distances from any of the three planes through the centers and parallel to the faces are less than $\alpha_2\delta_1/2$. There remain sixty-four cubes C_2 of side $\delta_2 = (1 - \alpha_2)\delta_1/2$. Continuing in this way, there are defined 8^n cubes C_n, of side $\delta_n = (1 - \alpha_n)\delta_{n-1}/2$, $n = 1, 2, 3, \cdots$. The set of points E common to all the cubes of the infinite set thus determined, is the *set of Cantor type corresponding to the sequence* $[\alpha_i]$.

3. *A Lemma on Capacity.* Using the notation of section 1, and denoting the capacity of E by $c(E)$, we have

THEOREM I. *If U is harmonic in T (and this includes regularity at infinity), and never negative, and if its lower and upper limits on the boundary of T lie between a and b, $0 < a \leqq b$, then*

$$m/b \leqq c(E) \leqq m/a,$$

where m is the mass producing U, as given by Gauss' integral.

Proof. Given ϵ, $0 < \epsilon < a$, we consider the domain $U < a - \epsilon$. This is an infinite domain, lying, with its boundary, in T. Its conductor potential is $U/(a - \epsilon)$, and hence the capacity of its boundary is $m/(a - \epsilon)$. But the capacity of a set is never greater than that of an including set,* and hence $c(E) \leqq m/(a - \epsilon)$. As $c(E)$ is independent of ϵ, the second inequality of the theorem is established.

For the first inequality, we note that U/b never exceeds 1 near the boundary of T, and that it is therefore less than 1 in T. Accordingly, the conductor potentials of the sequence of domains approximating to T all exceed U/b in these domains. Hence their limit, V, the conductor potential of T, is nowhere less than U/b. If $V - U/b$ vanishes at any point of T, it is identically 0, by Gauss' theorem of the arithmetic mean, and then $c(E) = m/b$. Otherwise, there is an equipotential surface, $V - U/b = k$, $k > 0$, enclosing E, on which the derivative of $V - U/b$ in the direction of the normal,

* See Kellogg, *Foundations of Potential Theory*, Berlin, 1929, p. 331. Here further properties of capacity are treated, and the method of sequences described, pp. 322-338.

(counted as positive in the sense in which it points into the infinite region bounded by the equipotential surface), is never positive, and somewhere negative. Gauss' integral, extended over this surface, yields

$$c(E) - m/b = -\frac{1}{4\pi}\int\int_S \frac{\partial}{\partial n}(V - U/b)\,dS > 0,$$

and the first inequality of the theorem is established.

4. *Inequalities between the Capacities of Certain Sets and Subsets.* Let C denote a closed cube of side δ, and C' any one of the eight cubes formed from C by discarding the points whose distances from any of the three planes through the center and parallel to the faces is less than $\delta\alpha/2$ $(0 < \alpha < 1)$. Let e denote any set of points in one of the cubes C', and E the set consisting of e and the seven congruent and symmetrically placed sets in the remaining cubes C'. We desire inequalities on $c(E)$ in terms of $c(e)$.

We consider first the sum U of the conductor potentials of the eight sets e, at the points of E. Let P be such a point. The conductor potential of the subset e to which P belongs, does not exceed 1. The value at P of the conductor potential of any of the other sets e does not exceed $c(e)$ divided by the distance between the two cubes C', one containing P, and the other the subset in question.* Hence

$$U(P) \leqq 1 + \frac{c(e)}{\alpha\delta}\left\{3 \times 1 + 3 \times \frac{1}{2^{1/2}} + \frac{1}{3^{1/2}}\right\} = 1 + \frac{c(e)\mu''}{\alpha\delta},$$

where μ'' is a number, about 5.701. As the mass producing U is $8c(e)$, it follows from theorem I that

$$c(E) \geqq 8c(e)/[1 + c(e)\mu''/\alpha\delta].$$

To obtain an upper bound for $c(E)$, we first replace each set e by an including set $\bar{e}$, bounded by one or more surfaces such that the Dirichlet problem is possible for the infinite complement of $\bar{e}$ and any continuous boundary values. This can be done in such a way that the sets $\bar{e}$ lie in the cubes C', for instance by enclosing the set e, with its limit points, in a finite number of spheres, and discarding the portions of the spheres outside of the corresponding cube C'. Distinguishing by a bar the quantities for the sets $\bar{e}$ from the corresponding quantities for the sets e, we may reason as follows. The value at a point P of $\bar{E}$ of the conductor potential of the set $\bar{e}$ to which P belongs, is 1, because the Dirichlet problem is possible for the infinite domain of the complement of $\bar{e}$. The value at P of each of the conductor

* See *Foundations of Potential Theory, l. c.*, p. 331, exercise 4.

potentials of the seven remaining sets $\bar{e}$ is not greater than $c(\bar{e})$ divided by the greatest distance from P to any of its points (or any greater distance). Hence

$$\bar{U}(P) \geqq 1 + c(\bar{e})\,\mu'/\delta,$$

where μ' is a number, about 5.026. It follows from theorem I that

$$c(\bar{E}) \leqq 8c(\bar{e})/[1 + c(\bar{e})\mu'/\delta].$$

In the limit, as the enveloping surfaces shrink down on the sets e, we have the same inequality for the sets E and e.

We shall find it convenient to use the inequalities derived in the form

$$(1) \qquad 1/8c(e) + \mu'/8\delta \leqq 1/c(E) \leqq 1/8c(e) + \mu''/8\alpha\delta.$$

Since $c(e) = 0$ implies $c(E) = 0$, and conversely, vanishing capacities need not be excluded from the inequalities in this form, provided they are understood in this sense.

5. *A Necessary Condition that a Set of Cantor Type Have Positive Capacity.* Let E denote a set of Cantor Type, and let γ_n denote the capacity of that portion of E in one of the cubes C_n, *i. e.* $\gamma_n = c(E \cdot C_n)$. The first of the inequalities (1) then yields

$$1/\gamma_n \geqq 1/8\gamma_{n+1} + \mu'/8\delta_n.$$

Replacing n successively by $0, 1, 2, \cdots n$, and adding the resulting inequalities, we have

$$1/\gamma_0 = 1/c(E) \geqq (\mu'/8)\,[1 + 1/8\delta_1 + 1/8^2\delta_2 + \cdots + 1/8^n\delta_n] + 1/8^{n+1}\gamma_{n+1},$$
$$(n = 1, 2, 3, \cdots).$$

The two terms on the right are never negative, and we therefore infer

THEOREM II. *Necessary conditions that the capacity of the Cantor set E be positive, are* (a) *that the series*

$$(2) \qquad \sum_1^\infty 1/8^n\delta_n = \sum_1^\infty 1/[4^n \prod_1^n (1 - \alpha_i)],$$

converge, and (b) *that $8^n\gamma_n$ have a positive lower bound.*

Incidentally, we remark that the series (2) will converge if $\limsup \alpha_n < 3/4$. It will diverge if $\liminf \alpha_n > 3/4$.

6. *A Sufficient Condition that a Set of Cantor Type Contain Regular Points.* From the criterion of Wiener for the regularity of points may be derived the following *:

* See Kellogg and Vasilesco, *loc. cit.*, p. 519.

The point P of E is regular or exceptional, according as the integral

$$\int^1 \frac{c(\rho)}{\rho^2}\, d\rho$$

is divergent or convergent, $c(\rho)$ being the capacity of the portion of E in the sphere of radius ρ about P. It will be convenient to know that in this integral criterion, the function $c(\rho)$ may be replaced by the capacity $C(\rho)$ of the portion of E in a cube of fixed orientation, of half-side ρ, and with center at P. But this fact is made evident by the inequalities

$$\int^1 \frac{C(\rho/3^{1/2})}{\rho^2}\, d\rho = \frac{1}{3^{1/2}} \int_{\epsilon/3^{1/2}}^{1/3^{1/2}} \frac{C(\rho)}{\rho^2}\, d\rho \leqq \int_\epsilon^1 \frac{c(\rho)}{\rho^2}\, d\rho \leqq \int_\epsilon^1 \frac{C(\rho)}{\rho^2}\, d\rho,$$

which show that the integrals involving $c(\rho)$ and $C(\rho)$ converge or diverge together.

If P be taken at a corner of the unit cube C, then $C(\delta_n) = \gamma_n$, and the second inequality (1) will give us information about the function $C(\rho)$. Applied to the subsets in C_n and C_{n+1}, it yields

$$1/\gamma_n \leqq 1/8\gamma_{n+1} + \mu''/8\alpha_{n+1}\delta_n.$$

Combining this inequality with those obtained from it by replacing n by $n+1$, $n+2, \cdots n+p-1$, we find

$$\frac{1}{\gamma_n} \leqq \frac{\mu''}{8}\left[\frac{1}{\alpha_{n+1}\delta_n} + \frac{1}{8\alpha_{n+2}\delta_{n+1}} + \cdots + \frac{1}{8^{p-1}\alpha_{n+p}\delta_{n+p-1}}\right] + \frac{1}{8^p\gamma_{n+p}}.$$

We should now like to allow p to become infinite, discarding the remainder term outside the bracket. But it is not clear that its limit is 0. The difficulty can be turned by replacing E, for the moment, by the set of *all* points in all the cubes C_{n+p}; if we then denote by $\gamma_{n,p}$ the capacity of the portion of this set in a cube C_n, we have $\gamma_n = \lim\limits_{p\to\infty} \gamma_{n,p}$, and $c(C_{n+p}) = K\delta_{n+p}$, where K is the capacity of the unit cube. The above inequality then becomes

$$\frac{1}{\gamma_{n,p}} \leqq \frac{\mu''}{8}\left[\frac{1}{\alpha_{n+1}\delta_n} + \frac{1}{8\alpha_{n+2}\delta_{n+1}} + \cdots + \frac{1}{8^{p-1}\alpha_{n+p}\delta_{n+p-1}}\right] + \frac{1}{8^p K\delta_{n+p}}.$$

If we confine ourselves to the case which alone interests us, namely that in which the capacity of E is positive, the series (2) converges, and the term outside the bracket approaches 0 as p becomes infinite. The result is

$$1/\gamma_n \leqq (\mu''/8)\, 8^n R_{n-1}, \tag{3}$$

where R_n is the remainder, after n terms, of the series

$$\sum_0^\infty 1/8^i\alpha_{i+1}\delta_i, \qquad (\delta_0 = 1). \tag{4}$$

Returning, with the inequality (3), to the integral criterion for regularity, we note that

$$\int_{\delta_{n+1}}^{\delta_n} [C(\rho)/\rho^2] d\rho \geqq \gamma_{n+1} \int_{\delta_{n+1}}^{\delta_n} d\rho/\rho^2 = \gamma_{n+1}(1/\delta_{n+1} - 1/\delta_n),$$

and hence that

$$(5) \qquad \int_0^1 [C(\rho)/\rho^2] d\rho \geqq \sum_1^\infty \gamma_{n+1}(1/\delta_{n+1} - 1/\delta_n)$$

$$\geqq (8/\mu'')\sum_1^\infty (1/8^{n+1}R_n)(1/\delta_{n+1} - 1/\delta_n) = (8/\mu'')\sum_1^\infty [1/8^{n+1}\delta_{n+1}R_n]\,[1 + \alpha_{n+1})/2].$$

As α_n lies between 0 and 1, we arrive at

THEOREM III. *A sufficient condition that the set E of Cantor type have a regular point, is that $8^n\delta_n$ become infinite with n, and that the series*

$$(6) \qquad \sum_1^n 1/8^{n+1}\delta_{n+1}R_n$$

diverge.

In order to apply this condition, we shall have need of a theorem on infinite series with positive terms, which is a special case of one due to Dini,* and which may be stated as follows:

If $\sum_1^\infty c_n$ is a convergent series of positive terms and if $r_{n-1} = c_n + c_{n+1} + \cdots$ denotes its remainders, then the series

$$\sum_1^\infty c_n/r_{n-1}$$

is divergent.

7. *Establishment of the Fundamental Lemma in Two General Cases.* We proceed to establish the lemma in the two cases (a) in which the terms of the sequence $[\alpha_i]$ have a positive lower bound, and (b) in which this sequence has the unique limit 0.

In the first case, on the hypothesis that E has positive capacity, the series (2) converges, by theorem II. Since the α_i have a positive lower bound, the series (4) also converges. It then follows from the theorem of Dini that the series

$$\sum_1^\infty 1/8^{n+1}\delta_{n+1}\alpha_{n+2}R_n$$

diverges, and from this follows the divergence of the series (6). The con-

* See Knopp, *Theorie und Andwendung der unendlichen Reihen*, Berlin 1922, p. 285, 2nd ed., (1924), p. 294, or Fort, *Infinite Series*, Oxford 1930, pp. 42-43.

vergence of (2) implies that $8^n\delta_n$ becomes infinite with n, and hence by theorem III, E has the regular point P.

The second case can be reduced to the first by the following device. We form a subset E' of E by omitting those points of E which are in those cubes C_2 which are not at the corners of C, omitting the points in those cubes C_4 which are not at the corners of cubes C_2, and so on. Those points of E are omitted which are in those cubes C_{2n+2} which are not at the corners of cubes C_{2n}. It is at once clear that E' is a set of Cantor type, and a little reckoning shows that the sequence $[\alpha_i']$, which characterizes it, is as follows

$$\begin{aligned}\alpha_1' &= \alpha_1 + (1-\alpha_1)(1+\alpha_2)/2\\ &= 1/2 + [\alpha_2 + \alpha_1(1-\alpha_2)]/2,\\ \alpha_2' &= 1/2 + [\alpha_4 + \alpha_3(1-\alpha_4)]/2,\\ &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ \alpha_n' &= 1/2 + [\alpha_{2n} + \alpha_{2n-1}(1-\alpha_{2n})]/2,\\ &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\end{aligned}$$

The expression for α_n' shows that $\lim \alpha_n' = 1/2$, and it follows from the remark following theorem II that the series (2) for E' converges. The same expression shows that $\alpha_n' > 1/2$, and it follows then from case (a) that the series (6) for E' diverges. Theorem III then shows that E' has regular points. Hence, E, which contains E', also has regular points.

We remark that the reasoning employed above shows that in the cases studied the vertices of all the cubes C_n are regular points. Thus the regular points, in these cases, are everywhere dense in E. If the fundamental lemma is true, this property is possessed by all reduced * sets of positive capacity.

The fundamental lemma is thus established for sets of Cantor type, except when the set $[\alpha_n]$ has two or more limit points, one of them 0. The series set up enable one to study further cases, but I have not settled them all. On the other hand, no cases have come to light in which the lemma fails.

8. *Two Remarks on Capacity in General.* We formulate the first remark as

THEOREM IV. *Any closed bounded set of positive Jordan outer content contains regular points.* On the other hand, the set of Cantor type formed with $\alpha_i = 1/2$ shows that a set may have 0 outer content, and still have regular points, by theorem III.

If the set is a plane set, we may understand the content referred to as plane content. Otherwise, we are to understand the three dimensional content. Regularity refers to behavior with respect to the Newtonian potential.

We give the proof of the theorem for the case of a plane set. Only

* **For the definition of reduced sets, see Kellogg and Vasilesco, *loc. cit.*, pp. 519-520.**

formal modifications are necessary for the case in which three dimensional content is involved. Suppose then, that E is a closed set, lying in the unit square, with the outer content $a > 0$. If we divide the square into four equal squares, the outer content of the portion of E in at least one of them will be not less than $a/4$. Call such a square S_1. If S_1 be similarly divided into four equal squares, the part of E in at least one of them will have an outer content not less than $a/4^2$. Continuing this process, we construct an infinite sequence of squares $S_1, S_2, S_3, \cdots$, such that the outer content of the part of E in S_n is at least $a/4^n$. If P is the limit of this sequence, then P is a regular point of E, as we now show.

Let γ_n denote the capacity of the portion E_n of E in S_n. Then E_n can be enclosed in a set of squares of a quadratic mesh so fine that the capacity of this set of squares differs arbitrarily little from γ_n, while the area A of the set of squares is not less than $a/4^n$. Let U denote the potential of a spread of unit surface density on these squares. Then U cannot exceed, at any point P, the potential at P of a spread of unit surface density on a circle of area A with center at P.* It follows that $U \leqq 2(\pi A)^{1/2}$. As the mass producing U is A, theorem I assures us that the capacity of the set of squares is not less than $\frac{1}{2}(A/\pi)^{1/2} \geqq \frac{1}{2^{n+1}}(a/\pi)^{1/2}$. Since this last expression is independent of the fineness of the mesh, it follows that it is also a lower bound for γ_n. Using this lower bound in the integral criterion of section 6, we see that P is indeed a regular point of E. In fact, since $C(\rho)$ is a monotonically increasing function of ρ, $C(\rho)/\rho$ has a positive lower bound, and the integral is hence divergent.

The second remark concerns the question of the possible topological character of the capacity of a set. Formulating it as broadly as possible, we ask: does a continuous one-to-one transformation, not only of the set E, but of the whole of space, necessarily carry E into a set E' which has positive capacity when, and only when, E has?

The sets of Cantor type furnish an immediate negative answer. In fact, let the set E be characterized by the sequence in which α_i has the constant value $3/5$, and E' by the sequence in which α_i has the constant value $4/5$. Then E has regular points, and hence positive capacity, by theorem III, while E' has 0 capacity, by theorem II. And there is no difficulty in setting up a continuous one-to-one transformation of space, carrying E into E'. These sets therefore constitute an example in point.

CAMBRIDGE, MASS.,
NOVEMBER 22, 1930.

---

* See *Foundations of Potential Theory*, *loc. cit.*, p. 149, lemma III (b).

NOTE ON FRACTIONAL OPERATORS AND THE THEORY OF COMPOSITION.

By Leonard M. Blumenthal.

In this paper the methods and notation of the theory of composition of the first kind, developed by V. Volterra and J. Pérès,* are applied to certain fractional operators.† The concept of the "finite part" of an integral developed almost simultaneously by J. Hadamard and R. D'Adhemar ‡ is employed to give a new definition of negative fractional orders of integration that is amenable to treatment as a very special case of composition. The symbolism already in use in the theory of composition is utilized with obvious advantage in the theory dealt with in this paper. Finally, definitions are given for the operators in the case of the complex variable, the concept of "finite part" being extended to contour integrals.

1. If a function $F(x,y)$ can be put in the form

$$F(x,y) = \frac{(y-x)^{\alpha-1}}{\Gamma(\alpha)}\,\phi(x,y),$$

where α is different from zero or a negative integer, and the function $\phi(x,y)$, finite and continuous, is not zero for $y=x$, then $F(x,y)$ is said to be of regular order α.

The resultant (Volterra product)

$$FG = \int_x^y F(x,\xi)\,G(\xi,y)\,d\xi$$

of two functions of regular orders α and β is, by a fundamental theorem, of order $\alpha+\beta$.

For $\alpha > 0$, the integral of order α has been defined by the expression §

* A systematic development of the theory is to be found in Volterra et Pérès, *Leçons sur la composition*, Paris (1924).

† For a historical sketch as well as a bibliography see "A Survey of Methods for the Inversion of Integrals of Volterra type," H. T. Davis, *Indiana University Studies*, No. 76, 77; also two papers by the same author, *American Journal of Mathematics*, Vol. 46 (1924), pp. 95-109; Vol. 49 (1927), pp. 123-142.

‡ J. Hadamard, *Annales de l'École Normale Supérieure* (1905), p. 122; R. D'Adhemar, *Exercices et leçons d'analyse*, Paris (1908), pp. 150, 180.

§ G. H. Hardy, "Notes on Some Points in the Integral Calculus," *Messenger of Mathematics*, Vol. 47 (1917-18), p. 145.

$${}_0F_\alpha[f(y)] = [1/\Gamma(\alpha)] \int_0^y (y-t)^{\alpha-1} f(t)\,dt,$$

where the function f is summable throughout any interval of positive values of y, $y=0$ included. Hardy, in the paper referred to below, shows that the expression exists for almost all values of y and is summable. It is to be noted that the lower limit of the integral is an essential part of the definition.

Now *

$$1^\alpha = (y-x)^{\alpha-1}/\Gamma(\alpha).$$

Hence we may write, putting $F = f(y)$,

$${}_xF_\alpha[f(y)] = f1^\alpha,$$

the lower limit x being a parameter which hereafter we shall omit exhibiting. The whole theory of integrals of positive order may thus be treated by means of the theory of composition. It is immediate, for example, that the operator $F_\alpha[f(y)]$ obeys the index law

$$F_\mu F_\nu[f(y)] = F_{\mu+\nu}[f(y)], \qquad \mu > 0,\ \nu > 0$$

for

$$F_\nu[f(y)] = f1^\nu,$$

and

$$F_\mu F_\nu[f(y)] = (f1^\nu)1^\mu.$$

But composition is associative, and hence

$$F_\mu F_\nu[f(y)] = f(1^\nu 1^\mu),$$

which, by means of the theorem cited above on the order of the resultant, yields

$$F_\mu F_\nu[f(y)] = f1^{\mu+\nu} = F_{\mu+\nu}[f(y)].$$

Again, making use of the concept of fractions of composition it is immediate that

$$\lim_{\nu\to 0} F_\nu[f(y)] = f(y).$$

2. When two functions are of negative orders, the integral defining their composition may not exist. It has been found, however, that all the theorems valid for ordinary composition hold if we agree, that whenever the integral FG does not exist, we are to take the *finite part* of the integral.† It is upon this that we base our definition of negative fractional order of

* It is usual to place an asterisk above the 1 to indicate the Volterra product.

† J. Pérès, "Sur la composition 1ère espèce: Les fonctions d'ordre quelconque et leur composition," *Rendiconti della Realle Accademia dei Lincei*, (1917), p. 45, 104.

integration. *Without introducing any new notation, it is understood that the finite part of all non-convergent integrals is to be taken.*

Let now $\alpha = -\eta$, $0 < \eta < 1$. We define the integral of negative order $-\eta$ by the expression

$$F_{-\eta}[f(y)] = 1/[\Gamma(-\eta)] \int_x^y (y-t)^{-\eta-1} f(t)\,dt.$$

3. *Finite part of* $\int_x^y (y-t)^{-\eta-1} f(t)\,dt$. Let the function f satisfy a Lipschitz condition throughout any interval of positive values of y, and consider the expression

$$I = \lim_{\epsilon \to 0} \left[\int_x^{y-\epsilon} (y-t)^{-\eta-1} f(t)\,dt + \phi(y-\epsilon)\epsilon^{-\eta} \right],$$

where

(1) $$f(y) = -\eta\phi(y).$$

THEOREM. *The limit I exists.*

We note first that I differs from

$$I' = \lim_{\epsilon \to 0} \left[\int_x^{y-\epsilon} \frac{f(t)-f(y)}{(y-t)^{1+\eta}}\,dt + \frac{\phi(y-\epsilon)-\phi(y)}{\epsilon^\eta} \right]$$

by a finite term $f(y)/\eta(y-x)^\eta$; for

$$I' = \lim_{\epsilon \to 0} \left[\int_x^y (y-t)^{-\eta-1} f(t)\,dt + \epsilon^{-\eta}\phi(y-\epsilon) - \epsilon^{-\eta} \left\{ \frac{f(y)}{\eta} + \phi(y) \right\} + \frac{f(y)}{\eta(y-x)^\eta} \right]$$

whence, using (1)

(2) $$I' = I + f(y)/\eta(y-x)^\eta.$$

Now the limit I' is readily seen to exist, for

$$I' \leqq \lim_{\epsilon \to 0} \left[\int_x^{y-\epsilon} \frac{|f(t)-f(y)|}{|(y-t)|^{1+\eta}}\,dt + \frac{|\phi(y-\epsilon)-\phi(y)|}{\epsilon^\eta} \right],$$

and using the Lipschitz condition

$$I' < k \lim_{\epsilon \to 0} \left[\int_x^{y-\epsilon} (y-t)^{-\eta}\,dt + \epsilon^{1-\eta} \right], \ k \text{ a finite constant,}$$

or

$$I' < [k/(1-\eta)](y-x)^{1-\eta}.$$

Whence, from (2) we infer that the limit I exists. We call I the *finite part* of the integral

$$\int_x^y (y-t)^{-\eta-1} f(t)\,dt,$$

and hence, by our definition

$$F_{-\eta}[f(y)] = \lim_{\epsilon\to 0}\left[\int_x^{y-\epsilon} \frac{(y-t)^{-\eta-1}f(t)}{\Gamma(-\eta)}\,dt + \frac{\phi(y-\epsilon)}{\Gamma(-\eta)\epsilon^{\eta}}\right],$$

If the function f is differentiable, this expression can be put in a more useful form. First we write $F_{-\eta}[f(y)]$ in the equivalent form

$$(3)\quad F_{-\eta}[f(y)] = \lim_{y_1\to y}\left[\int_x^{y_1} \frac{(y-t)^{-\eta-1}f(t)}{\Gamma(-\eta)}\,dt + \frac{\phi(y_1)}{\Gamma(-\eta)(y-y_1)^{\eta}}\right], \quad y_1 < y.$$

Now

$$\frac{d}{dt}\frac{\phi(t)}{(y-t)^{\eta}} = \frac{\eta\phi(t) + \phi'(t)(y-t)}{(y-t)^{1+\eta}}$$

whence

$$\frac{\phi(y_1)}{(y-y_1)^{\eta}} = \int_x^{y_1} \frac{\eta\phi(t)+\phi'(t)(y-t)}{(y-t)^{1+\eta}}\,dt + \frac{\phi(x)}{(y-x)^{\eta}}.$$

Substituting in (3), and taking account of (1), we have

$$(4)\quad F_{-\eta}[f(y)] = \frac{\phi(x)}{\Gamma(-\eta)(y-x)^{\eta}} + \int_x^y \frac{\phi'(t)}{\Gamma(-\eta)(y-t)^{\eta}}\,dt = \frac{1}{\Gamma(-\eta)}\frac{d}{dy}\int_x^y \frac{\phi(t)}{(y-t)^{\eta}}\,dt.$$

In general, if f is differentiable $(p-1)$-times and its $(p-1)$-st derivative satisfies a Lipschitz condition,

$$(5)\quad \int_x^y (y-t)^{-p-\mu} f(t)\,dt = \int_x^y (y-t)^{-p-\mu} f_1(t)\,dt + \int_x^y \frac{f(t)-f_1(t)}{(y-t)^{p+\mu}}\,dt, \quad 0<\mu<1$$

where

$$f_1(t) = f(y) - f'(y)(y-t) + f''(y)(y-t)^2/2! + \cdots + (-1)^{p-1} f^{(p-1)}(y)(y-t)^{p-1}/(p-1)!.$$

Now

$$\int_x^y (y-t)^{-p-\mu}\,dt = 11^{-p-\mu+1}\Gamma(-p-\mu+1) = (y-x)^{1-p-\mu}/(1-p-\mu).$$

The first member of the right hand side of (5) is a sum of terms of the form

$$[f^{(i)}(y)/i!]\int_x^y (y-t)^{-p-\mu+i}\,dt$$

and hence may be written

$$\sum_{i=0}^{p-1} (-1)^i [f^{(i)}(y)/i!] \cdot (y-x)^{1+i-p-\mu}/(1+i-p-\mu).$$

Whence

$$\int_x^y \frac{f(t)}{(y-t)^{p+\mu}}\,dt = \sum_{i=0}^{p-1} (-1)^i \frac{f^{(i)}(y)\,(y-x)^{1+i-p-\mu}}{(1+i-p-\mu)i!} + \int_x^y \frac{f(t)-f_1(t)}{(y-t)^{p+\mu}}\,dt.$$

3. A few examples of functions obtained by the use of the definitions used in this paper are given. In this table we have set the parameter x equal to zero. The functions are obtained with a minimum of calculation.

$$F_{-\eta}[k] = k \cdot 11^{-\eta} = k \cdot 1^{1-\eta} = [k/\Gamma(1-\eta)]y^{-\eta},$$

$$F_{-\eta}[y^k] = 1^{k+1}1^{-\eta}\Gamma(1+k) = [\Gamma(1+k)/\Gamma(k-\eta+1)]y^{k-\eta},$$

$$F_{-\eta}[e^y] = \sum_{i=1}^{\infty} 1^i 1^{-\eta} = \sum_{i=1}^{\infty} y^{i-\eta-1}/\Gamma(i-\eta), \text{ also, from (4),}$$

$$F_{-\eta}[e^y] = \frac{y^{-\eta}}{\Gamma(1-\eta)} + e^y \int_0^y \frac{e^{-t}t^{-\eta}}{\Gamma(1-\eta)}\,dt,$$

$$F_{-\eta}[\sin y] = \sum_{p=1}^{\infty} (-1)^{p-1} 1^{2p} 1^{-\eta} = \sum_{p=1}^{\infty} (-1)^{p-1}[y^{2p-\eta-1}/\Gamma(2p-\eta)].$$

For $f(t)$ representable in a series

$$f(t) = a_0 + a_1 t + a_2 t^2/2! + \cdots + a_n t^n/n! + \cdots,$$

convergent for $0 \leqq t \leqq y$, we have

$$F_{-\eta}[f(y)] = \sum_{i=0}^{\infty} [a_i/\Gamma(i-\eta)]y^{i-\eta-1}.$$

4. *Differentiation.* The finite part of the integral

$$\int_x^y (y-t)^{-p-\mu} f(t)\,dt$$

admits directly of differentiation with respect to y. It is performed by differentiating under the integral sign as though the integrand were continuous at y and ignoring the fact that the upper limit of the integral is a function of the parameter. Thus

$$F_{-\eta}[f(y)] = [1/\Gamma(-\eta)] \int_x^y (y-t)^{-\eta-1} f(t)\,dt;$$

$$dF_{-\eta}/dy = [(-1-\eta)/\Gamma(-\eta)] \int_x^y (y-t)^{-\eta-2} f(t)\,dt,$$

and since

$$[(-1-\eta)/\Gamma(-\eta)] = [1/\Gamma(-1-\eta)],$$

we have

$$dF_{-\eta}/dy = [1/\Gamma(-1-\eta)] \int_x^y (y-t)^{-\eta-2} f(t)\,dt = F_{-\eta-1}[f(y)].$$

It is readily seen that the operator satisfies the condition of linearity and the index law. The latter is immediate upon using the classical properties of the finite part, or may be even more readily seen by use of the theorems valid for composition of functions of negative orders. Thus

$$F_{-\mu}F_{-\nu}[f(y)] = F_{-\mu}(f1^{-\nu}) = (f1^{-\nu})1^{-\mu} = f(1^{-\nu}1^{-\mu}) = f1^{-\mu-\nu} = F_{-\mu-\nu}[f(y)].$$

Example. Let us apply our definition of fractional operator to obtain the solution of Abel's equation:

$$\phi(y) = \int_0^y (y-t)^{-\eta}u(t)\,dt = u1^{1-\eta}\Gamma(1-\eta), \quad 0 < \eta < 1,$$

where the function $\phi(y)$ satisfies a Lipschitz condition (and is therefore absolutely continuous) and is not necessarily zero for $y = 0$. Operating with the symbol $1^{\eta-1}$, we have

$$u(y) = [1/\Gamma(1-\eta)]\,\phi 1^{\eta-1} = [1/\Gamma(1-\eta)]\,F_{\eta-1}[\phi(y)],$$

where, in accordance with our agreement we are to take the *finite part* of the *non-convergent integral* appearing in the right-hand member. But by our definition, the finite part of $[1/\Gamma(1-\eta)]\,F_{\eta-1}[\phi(y)]$ is readily given, almost everywhere, by the expression

$$\frac{\sin\pi\eta}{\pi}\left[\frac{\phi(0)}{y^{1-\eta}} + \int_0^y \frac{\phi'(t)}{(y-t)^{1-\eta}}\,dt\right].$$

Hence

$$u(y) = \frac{\sin\pi\eta}{\pi}\left[\frac{\phi(0)}{y^{1-\eta}} + \int_0^y \frac{\phi'(t)}{(y-t)^{1-\eta}}\,dt\right]$$

almost everywhere. But L. Tonelli has shown that the function $u(y)$ determined for almost all values of t by this expression satisfies Abel's equation for all values of the variable y. Thus, substituting, we have

$$\begin{aligned}\int_0^y \frac{u(t)}{(y-t)^{\eta}}\,dt &= \frac{\sin\pi\eta}{\pi}\,\phi(0)\int_0^y \frac{dt}{t^{1-\eta}(y-t)^{\eta}} \\ &\quad + \frac{\sin\pi\eta}{\pi}\int_0^y \frac{dt}{(y-t)^{\eta}}\int_0^t \frac{\phi'(z)}{(t-z)^{1-\eta}}\,dz \\ &= \phi(0) + \frac{\sin\pi\eta}{\pi}\int_0^y \phi'(z)\int_z^y (y-t)^{-\eta}(t-z)^{\eta-1}\,dt\,dz \\ &= \phi(0) + \int_0^y \phi'(z)\,dz \\ &= \phi(y).\end{aligned}$$

* L. Tonelli, "Su un Problema di Abel," *Mathematische Annalen*, Vol. 99 (1928), p. 191.

5. *The Fractional Operator in the Complex Domain.* If n is a positive integer and $f(t)$ is holomorphic within a region S whose boundary C consists of a finite number of regular curves, and if the function f is continuous along C, then the n-th derivative of f at the point z lying within C is given by

$$f^{(n)}(z)=(n!/2\pi i)\int_C (t-z)^{-n-1}f(t)\,dt.$$

Consider the expression

$$(6)\quad f^{-y}(z)=[\Gamma(1-y)/2\pi i]\int_C f(t)\exp\,[(y-1)\{\log(t-z)+2k\pi i\}]\,dt,$$

where k is an integer and the point z is a point of S or on C. We prove the following theorem:

THEOREM. $f^{-p}(z)=\lim_{y\to p} f^{-y}(z)$, *$p$, a positive integer, exists and differs from the iterated integral of order p by a polynomial $\phi(z)$ such that $f^{-p}(z)-\phi(z)$ is zero at the point $z=a$, together with its $(p-1)$ derivatives, where a is a point within or on C.*

For

$$\lim_{y\to p} f^{-y}(z)=\lim_{y\to p}\frac{\Gamma(1+p-y)}{(1-y)(2-y)\cdots(p-1-y)}\cdot\frac{(1/2\pi i)\int_C f(t)\exp\,[y-1)\{\log(t-z)+2k\pi i\}]\,dt}{y-p}$$

$$=\frac{(1/2\pi i)}{(-1)^p(p-1)!}\int_C f(t)(t-z)^{p-1}\{\log(t-z)+2k\pi i\}dt,$$

which is independent of k.

Thus

$$f^{-p}(z)=\lim_{y\to p} f^{-y}(z)=[(1/2\pi i)/(-1)^p(p-1)!]\int_C f(t)(t-z)^{p-1}\log(t-z)\,dt,$$

of which the p-th derivative is $f(z)$. Whence

$$f^{-p}(z)-\phi(z)=\int_a^z dt_1\int_a^{t_1}dt_2\cdots\int_a^{t_{p-1}}f(t_p)\,dt_p,\ a \text{ in } S.$$

The operator $f^{-y}(z)$ defined in (6) is a mixed linear functional. For an integral index the functional is monodromic with monodromic indicatrix *

$$v(z,t)=-n!/(t-z)^{n+1}.$$

---

* L. Fantappiè, "I Funzionali Analitici," *Sem. Matematico della Facoltà di Scienze*, Università di Roma, Rendiconti, 1925-26, ser. 2a, Vol. 14.

For a fractional index the operator is a mixed linear polydromic functional with polydromic indicatrix. Now it is natural to define the positive fractional order of integration in the complex domain by the expression

$$f_0^{-\mu}(z) = [\Gamma(1-\mu)/2\pi i] \int_c (t-z)^{-1+\mu} f(t)\,dt, \quad 0 < \mu < 1$$

where the curve C has the origin as its initial and final point and incloses the branch point z, and an assigned branch of the multiform integrand is holomorphic in the region bounded by C. The choice of the curve C to pass through the origin is an essential part of the definition, since integration along two closed curves around z not both starting and ending at the same point will give different values to the operator.

6. *Development of* $f_0^{-\mu}(z)$. The function $f(t)$ being expressible in the form $f(t) = \sum_{n=0}^{\infty} a_n t^n$, the series $(t-z)^{\mu-1} \sum_{n=0}^{\infty} a_n t^n$ can be integrated termwise, and we have

$$f_0^{-\mu}(z) = [\Gamma(1-\mu)/2\pi i] \sum_{n=0}^{\infty} a_n \int_c t^n (t-z)^{\mu-1}\,dt.$$

Now for C we chose a loop from the origin around the branch-point z and back to the origin. Whence

$$\int_c t^n (t-z)^{\mu-1}\,dt = (1-e^{2\pi i\mu}) \int_0^z t^n (t-z)^{\mu-1}\,dt,$$

where the integral may be considered taken along the straight line from the origin to the point z. If in this integral we put $t = \xi z$, we obtain

$$\int_c t^n (t-z)^{\mu-1}\,dt = (1-e^{2\pi i\mu})(-1)^{\mu-1} z^n z^\mu \int_0^1 \xi^n (1-\xi)^{\mu-1}\,d\xi.$$

Whence

$$f_0^{-\mu}(z) = (1/2\pi i)\Gamma(1-\mu)(1-e^{2\pi i\mu})e^{-\pi i(\mu-1)} z^\mu \sum_{n=0}^{\infty} a_n \beta(n+1,\mu) z^n$$
$$= z^\mu \sum_{n=0}^{\infty} \frac{a_n \Gamma(n+1)}{\Gamma(n+\mu+1)} z^n.$$

Thus this definition of the positive fractional order of integration in the complex domain as a contour or loop integral is seen to yield the same Taylor development as the definition for this operator adopted by Hadamard.* The

* The Taylor development for the Hadamard definition is given in Mandelbrojt, "Modern Researches on the Singularities of Functions Defined by Taylor's Series," *The Rice Institute Pamphlet*, Vol. 14 (1927), No. 4, p. 291.

contour integral, however, fits into the theory of composition in the complex field and serves to make the theory of this operator a part of the already extensively developed theory of complex functionals of V. Volterra and L. Fantappiè.*

In defining the negative fractional order of integration we use a device entirely similar to that of the finite part of an integral. Some such device is essential, since in integrating a function $t^n(t-z)^{-\mu-1}$ with $0<\mu<1$ along C, a loop from the origin around the point z and back to the origin, the integral around the circle with z as center *does not approach zero* with the radius of the circle. We write, therefore

$$ {}_0F_{-\mu}[f]=(1/2\pi i)\Gamma(1+\mu)\overline{\int_c (t-z)^{-\mu-1}f(t)\,dt}, $$

where, by definition

$$ \overline{\int_c (t-z)^{-\mu-1}f(t)\,dt}=(1-e^{-2\pi i\mu})\int_0^z (t-z)^{-\mu-1}f(t)\,dt, $$

with the *finite part* of the integral in the second member being taken. Thus

$$ {}_0F_{-\mu}[f]=(1/2\pi i)\Gamma(1+\mu)(1-e^{-2\pi i\mu})e^{\pi i(\mu+1)}\int_0^z (z-t)^{-\mu-1}f(t)\,dt $$

$$ =\frac{1}{\Gamma(-\mu)}\int_0^z (z-t)^{-\mu-1}f(t)\,dt=\frac{1}{\Gamma(1-\mu)}\left[\frac{f(0)}{z^\mu}+\int_0^z \frac{f'(t)}{(z-t)^\mu}\,dt\right] $$

7. *The Index Law for* $f_0^{-\mu}(z)$.

THEOREM. $f_0^{-\nu}f_0^{-\mu}(z)=f_0^{-\mu-\nu}(z),\quad 0<\mu,\ \nu<1,\quad \mu+\nu\neq 1.$

By definition we have

$$ f_0^{-\nu}f_0^{-\mu}(z)=[\Gamma(1-\nu)\Gamma(1-\mu)/(2\pi i)^2]\int_{C_1}(t-z)^{-1+\nu}dt\int_{C_2}(u-t)^{-1+\mu}f(u)\,du, $$

where both C_1 and C_2 have their initial and end-points at the origin and inclose the points z and t respectively.

Now choosing C_2 as a loop from the origin around the point t in the positive direction and back to the origin, we have

$$ \int_{C_2}(u-t)^{\mu-1}f(u)\,du=(1-e^{2\pi i\mu})\int_0^t (u-t)^{\mu-1}f(u)\,du. $$

Choosing C_1 as a loop around the point z, and applying Dirichlet's principle, we obtain

* L. Fantappiè, *loc. cit.*

$$f_0^{-\nu} f_0^{-\mu}(z) = (\tfrac{1}{2}\pi i)^2 \Gamma(1-\nu)\Gamma(1-\mu)(1-e^{2\pi i\mu})(1-e^{2\pi i\nu}) \int_0^z f(u)\,du \int_u^z (t-z)^{\nu-1}(u-t)^{\mu-1}dt.$$

A change of variable in the second integral enables us to write

$$f_0^{-\nu} f_0^{-\mu}(z) = \frac{-(1-e^{2\pi i\nu})(1-e^{2\pi i\mu})\Gamma(1-\nu)\Gamma(1-\mu)\Gamma(\mu)\Gamma(\nu)}{(2\pi i)^2\Gamma(\mu+\nu)} \int_0^z (u-z)^{\mu+\nu-1} f(u)\,du.$$

Writing the expression as a loop integral around z the above expression becomes, after some reductions,

$$f_0^{-\nu} f_0^{-\mu}(z) = \frac{\Gamma(1-\mu-\nu)}{2\pi i} \int_{c_1} (u-z)^{\mu+\nu-1} f(u)\,du,$$

and hence

$$f_0^{-\nu} f_0^{-\mu}(z) = f_0^{-\mu-\nu}(z).$$

THE RICE INSTITUTE.

POTENTIALS OF GENERAL MASSES IN SINGLE AND DOUBLE LAYERS. THE RELATIVE BOUNDARY VALUE PROBLEMS.

By G. C. Evans and E. R. C. Miles.

1. *Introduction.* The potential due to the most general distribution of finite positive and negative mass deposited in a single layer on a closed surface S may be written in the form

$$(1) \qquad v(M) = \int_S \frac{1}{MP}\, d\mu(e_P)$$

where the mass function $\mu(e)$ is a completely additive function of point sets e on S. The most general distribution of mass in a double layer on S yields similarly the potential

$$(2) \qquad u(M) = \int_S \frac{\cos(MP, n_P)}{MP^2}\, d\nu(e_P)$$

where $\nu(e)$ is likewise a completely additive function; here n_P denotes the direction of the interior normal to S at P. In fact, for all the closed surfaces to be discussed the direction n is that of the interior normal to the surface, whether or not it may be interior to the region in question.

In terms of these potentials, by means of Stieltjes integral equations, one can solve generalized boundary value problems of the first and second kinds. The first boundary value problem is solved by (2) when the limiting values are given of the quantity $\int u\, d\omega$, extended over an arbitrary portion ω of S', which is a surface neighboring S, as S' approaches S. In the second boundary value problem, limiting values of the flux $\int dv/dn\, d\omega$ are similarly given, and the problem is solved in terms of (1). Special cases of these problems are the Dirichlet and Neumann problems, respectively, with boundary values summable on S.*

2. *Differential geometry of S and of its neighborhood.* We assume that S is a simple closed surface with a tangent plane at every point, whose orientation changes continuously with respect to displacement of its point of tan-

* A summary of this paper appeared in the *Proceedings of the National Academy of Sciences*, Vol. 15 (February, 1929), pp. 102-108. Subsequently, the special case of the Neumann problem, the $\mu(e)$ being subject to the restriction of absolute continuity, was treated by M. Gunther, "Sur une application des intégrales de Stieltjes au problème de Neumann," *Comptes Rendus de l'Academie des Sciences*, Vol. 189 (September, 1929), pp. 447-450.

gency on S. A further restriction which is of importance for the potentials (1) and (2) is that there shall be a constant Γ such that

$$(2.1) \qquad \int_S \frac{|\cos(MP, n_P)|}{MP^2} \, d\omega_P < \Gamma$$

irrespective of the position of the point M, and

$$(2.2) \qquad \int_S \frac{|\cos(QP, n_Q)|}{QP^2} \, d\omega_P < \Gamma$$

irrespective of the position of the point Q on S, $d\omega$ being the element of surface area of S. A general type of such surface is furnished by the following theorem.*

LEMMA. *Let $QP = s$ be the arc length of the curve of section of S made by the plane which is determined by n_Q and P, and $f(s)$ a positive, continuous, non-decreasing function of s such that $\int_0^s f(s)/s\, ds$ is a convergent integral. If there is a number δ' such that the inequality*

$$(2.3) \qquad |\sphericalangle(n_Q, n_P)| \leq f(s)$$

holds for all Q, P on S for which $s \leq \delta'$, then there is a constant Γ such that (2.1) and (2.2) are satisfied.

There is no loss of generality if we take δ' small enough so that $f(s) < 1$, $s \leq \delta'$.

We take ρ, θ, z cylindrical coördinates of S in the neighborhood of Q, where ρ, θ are polar coördinates, referred to Q, in the tangent plane at Q. Let $\omega(\delta, Q)$ denote the portion of S containing Q which is bounded by the curve on S determined by $\rho = \delta$.

For a point P_1 in $\omega(\delta, Q)$, where δ is small enough, we have $\rho_1 > \frac{1}{2} s_1$. In fact, since $s = \int_0^\rho \sec(s, \rho)\, d\rho$ and $\sec(s, \rho) \leq \sec(n_Q, n_s)$, where n_s is the normal to S at the point whose parameter is s, this relation holds for $s_1 \leq \delta'$. If, then, we let $\delta = \delta'/2$, $g(\rho) = f(2\rho)$, we have $f(s) \leq g(\rho)$ for $\rho \leq \delta$, where $g(\rho)$ is a continuous non-decreasing function of ρ such that

$$\int_0^\delta \frac{g(\rho)}{\rho} \, d\rho = m(\delta)$$

converges. Then (2.3) may be rewritten

* Curves in the plane that satisfy the relation analogous to (2.1) have been studied as "curves of class Γ." See G. C. Evans, "Fundamental Points of Potential Theory," *The Rice Institute Pamphlet*, Vol. 7 (1920), No. 4, pp. 252-329. See p. 261.

$$(2.4) \qquad | \sphericalangle (M_Q, M_P) | \leq g(\rho), \qquad (\rho \leq \delta = \delta'/2),$$

and we have also the relations

$$(2.5) \qquad \left.\begin{aligned} z &\leq 2\rho g(\rho) \\ | \cos(QP, n_Q) | &\leq 2g(\rho) \\ | \cos(n_Q, n_P) | &> 1/2 \\ | d\omega | &< 2\rho d\rho d\theta \end{aligned}\right\} (\rho \leq \delta).$$

With respect to the integral (2.2), then, we have

$$(2.6) \qquad \int_{\omega(\delta,Q)} \frac{| \cos(QP, n_Q) |}{QP^2} d\omega_P < 8\pi m(\delta).$$

But the point Q is distant by a not zero amount Δ from the set of points comprising $S - \omega(\delta, Q)$. Hence

$$\int_S \frac{| \cos(QP, n_Q) |}{QP^2} d\omega_P = \int_\omega + \int_{S-\omega} < 8\pi m(\delta) + (1/\Delta^2) \text{meas. } S,$$

and the theorem is proved with reference to the integral (2.2).

The integral (2.1) may be considered first in the particular case where M is a point Q of S:

$$I = \int_S \frac{| \cos(QP, n_P) |}{QP^2} d\omega_P = I_\delta + I_\delta',$$

where I_δ is extended over $\omega(\delta, Q)$, and $I_\delta' \leq \text{meas. } S/\Delta^2$.

If we let $\psi = \sphericalangle(QP, n_P) - \sphericalangle(QP, n_Q)$, we shall have $| \psi | \leq g(\rho)$, and therefore

$$| \cos(QP, n_P) | \leq | \cos(QP, n_Q) | + g(\rho).$$

Hence $I_\delta < 12\pi m(\delta)$, so that

$$(2.7) \qquad I < 12\pi m(\delta) + (1/\Delta^2) \text{meas. } S.$$

For the general case, M not on S, let $\Delta(M)$ denote the distance of M from S, and $\Gamma(M)$ the corresponding value of the integral (2.1). Then

$$\Gamma(M) \leq \text{meas. } S/[\Delta(M)]^2.$$

If the $\Gamma(M)$ are not bounded, for all M, there will be a sequence of points $\{M_i\}$, with $\lim \Gamma(M_i) = \infty$. Then, since $\lim \Delta(M_i) = 0$, the points M_i will have a limit point Q on S. Without loss of generality we may suppose (by taking subsequences if necessary) that Q is the unique limit of the sequence.

On account of the continuity in the orientation of the normal, we know

that there is at least one normal to S in the neighborhood of Q which passes through M_k.* We may therefore, writing Q_k for the foot of one such normal, suppose that $\lim Q_k = Q$, and accordingly $\lim \overline{M_kQ_k} = 0$. We have

$$\Gamma(M_k) = \Gamma_\delta(M_k) + \Gamma_\delta'(M_k), \tag{2.8}$$

where the integrations are extended over $\omega(\delta, Q_k)$ and $S - \omega(\delta, Q_k)$, respectively. The second term on the right is evidently bounded, $< K$, for all M_k with k sufficiently large.

As for $\Gamma_\delta(M_k)$, we may write

$$\sphericalangle(M_kP, n_P) = \sphericalangle(M_kP, n_{Q_k}) + \psi',$$

where

$$|\psi'| \leq |\sphericalangle(n_{Q_k}, n_P)| \leq g(\rho), \qquad M_kP \geq \rho,$$

ρ being measured in the tangent plane at Q_k. Hence as before

$$\Gamma_\delta(M_k) \leq \int_{\omega(\delta, Q_k)} \frac{|\cos(M_kP, n_Q)|}{M_kP^2}\, d\omega_P + 4\pi m(\delta).$$

But if we write $\sphericalangle(M_kP, n_Q) = \sphericalangle(M_kP', n_Q) + \psi''$, where P' is the projection of P on the tangent plane at Q_k, we have

$$|\cos(M_kP, n_Q)| \leq |\cos(M_kP', n_{Q_k})| + |\sin\psi''|,$$
$$\sin\psi'' \leq z/\rho.$$

Hence, writing ω' for the projection of $\omega(\delta, Q_k)$ on the tangent plane,

$$\Gamma_\delta(M_k) < 2\int_{\omega'} \frac{|\cos(M_kP', n_{Q_k})|}{M_kP^2}\, d\omega' + 2\int_{\omega'} (z/\rho^3)\, d\omega' + 4\pi m(\delta).$$

But

$$\left| \frac{1}{M_kP^2} - \frac{1}{M_kP'^2} \right| = \frac{|M_kP + M_kP'|\;|M_kP' - M_kP|}{M_kP^2 M_kP'^2} \leq (2z/\rho^3),$$

and therefore

$$\begin{aligned}\Gamma_\delta(M_k) &< 2\int_{\omega'} \frac{|\cos(M_kP', n_{Q_k})|}{M_kP'^2}\, d\omega' + 4\int_{\omega'} (z/\rho^3)\, d\omega' + 4\pi m(\delta)\\ &< 4\pi + 8\pi m(\delta) + 4\pi m(\delta).\end{aligned}$$

Hence, referring to (2.8), we see that $\Gamma(M_k)$ is bounded, contrary to our assumption.

It follows that $\Gamma(M)$ is bounded for all M not on S, and since by (2.7) it is bounded for all M on S, inequality (2.1) is established.

---

* In fact we may draw a small sphere with center Q and then choose M_k close enough to Q so that it is nearer to Q than to any point of the surface of the sphere. There will then be a shortest distance from M_k to points of S within the sphere, and this distance will fix a normal from M_k to S whose foot is the desired point Q_k.

A particular form of $f(s)$ which satisfies the conditions of the lemma is $f(s) = C s^{\alpha}$, $0 < \alpha$. *Hereafter it will be assumed, except in* § 6, *that* $f(s) < Cs$, *that is*

$$(2.9) \qquad |\sphericalangle(n_Q, n_P)| < Cs < 2C\rho, \qquad (\rho \leq \delta = \delta'/2).$$

Let now S' be a surface neighboring to S, simple and closed, with a tangent plane at every point, whose orientation changes continuously with respect to displacement of its point of tangency on S'.* Let it be defined by normal displacement of the point P of S in amount $n(P)$, $|n(P)| \leq \tau$, where $n(P)$ is a continuous non-vanishing function of P. If τ is small enough there will be a one to one correspondence between the points of S' and those of S. In fact, we see first that if P is a point of $\omega(\delta, Q)$ and n_P and n_Q intersect at M, we cannot have MP and $MQ < 1/C$. We refer S momentarily to rectangular coördinates at Q, the z-axis being in the direction of n_Q through M, and consider the plane section determined by n_Q and P. We take x along the intersection of this plane with the tangent plane at Q, in the direction of P, and assume that n_P and n_Q intersect at M.

For $\rho < \delta$, so that $s < \delta'$, we have the angle between n_Q and the normal to s at a point P' of s, in the plane section, $\leq$ the angle n_Q, n_P, and therefore $< Cs$, by (2.9). Hence

$$x = \int_0^s \cos(x, s)\, ds > \int_0^s \cos Cs\, ds = (1/C) \sin Cs$$

$$z = \int_0^s \sin(x, s)\, ds, \qquad |z| < \int_0^s \sin Cs\, ds = (1/C)(1 - \cos Cs).$$

Now for P in $\omega(\delta, Q)$ the angle (x, s) is less than 1 in numerical value, and accordingly $z_M > |z|$. Hence

$$|z_M - z| > |z_M - (1/C)(1 - \cos Cs)|$$

and

$$MP^2 = (z_M - z)^2 + x^2$$
$$MP^2 > z_M^2 + (2/C)[(1/C) - z_M](1 - \cos Cs).$$

From this inequality it follows that $MP > MQ$ if $z_M = MQ < 1/C$.

But similarly taking n_P as the z-axis, if $MP < 1/C$, we have $MQ > MP$. Hence both MP and MQ cannot be $< 1/C$.

Moreover the distance Δ from Q to the set of points $S - \omega(\delta, Q)$ has a lower bound Δ_1, for all Q, where $\Delta_1 > 0$. Hence if τ is less than $\Delta_1/2$

* The requirements on S' are not stated with sufficient precision in Evans and Miles, *loc. cit.*

no normal to a point of $S - \omega(\delta, Q)$ can cut n_Q in such a way that both normals are in length $< \tau$. Consequently if $\tau < \Delta_1$ and $< 1/C$ at the same time, the correspondence is one-one.

The orientation of the tangent plane to S' has been defined, according to hypothesis, by functions which are continuous at every point of S', and hence are uniformly continuous. The projection on the tangent plane at M of any element of arc ds on S', through R in the neighborhood of M is at least as great as $ds \cdot \cos(n_M, n_R)$; hence there is a small neighborhood of M of radius greater than some constant, independent of M, such that the projection on the tangent plane at M of any rectifiable arc s on S', in this neighborhood, will be a rectifiable arc on the tangent plane of length $\geq s/2$.

For the analysis of the boundary value problem we consider a family of surfaces $S' = S_\tau'$ where $| n(P) | < \tau$, and τ approaches zero. In order to complete the description of the relation of the family S_τ' to S we assume that there are positive constants C, α, δ'' such that if n_M denotes the normal to S' at a point M on S', and n_Q the normal to S at Q on S, the inequality

$$(2.10) \qquad | \sphericalangle (n_M, n_Q) | < C\overline{MQ}^\alpha$$

will be valid, whenever $MQ < \delta''$, irrespective of the particular surface S', of the family, involved.* As in the case of δ' and (2.9), the quantity δ'' is assumed to be small enough so that the angle described in (2.10) is < 1. It follows then that the projection on the tangent plane to S at Q of an element of arc ds' at M on S' is $\geq ds'/2$ if $MQ < \delta''$.

In particular, the requirements for S_τ' are satisfied if S_τ' is parallel to S internally or externally, and in its neighborhood, that is, if $n(P) = \tau$. Also, if S' is a surface with continuous curvatures, and if $n_0(P)$ is a continuous non-vanishing function on S, in absolute value ≤ 1, with continuous derivatives with respect to displacement on S, the family of surfaces S_τ', where $n(P) = \tau n_0(P)$, satisfies the above requirements.

A *regular curve* on S or on S', for the purposes of this article, will be a simple closed curve on the surface, made of a finite number of arcs, each with a continuously turning tangent, the two branches which come together making a not zero angle with each other. Since any part of such a curve, contained in a neighborhood of diameter sufficiently small, say equal to some constant δ''', when projected on the tangent plane at any point of this neighborhood, is itself a portion of a regular curve, and the projection of the

* This condition, rather than a slightly less restrictive one, is introduced in conformity to (2.9). The condition (2.9) was introduced in order to generate a one to one correspondence between the points of S' and those of S.

neighborhood is itself of diameter $\geq \frac{1}{2}\delta'''$, we see that any regular curve on S or on S' can be subdivided into regular cells of diameter uniformly small. Vice versa, given a small regular arc in the tangent plane, there corresponds to it a small regular arc on the surface. A simple closed curve w on the surface may then be described as a regular curve if it can be subdivided into a finite number of cells of diameter arbitrarily small, each of which has as its image, on the tangent plane at one of its points, a closed regular curve.

We define a *regular net* G on the surface, as a system of lattices G_n, corresponding to positive numbers δ_n successively decreasing and approaching zero as a limit. Each lattice represents a partition of the preceding lattice into a finite number of cells w_{in}, each of which is a regular closed curve of diameter $< \delta_n$, and G_0 is the surface itself. In a similar manner we may define a net G for a regular closed curve w on the surface, with reference to its interior region ω, and given such a net, we may extend it through the complement of ω so that it becomes a regular net for the surface itself.*

3. *Stieltjes Integrals.* We may form the Stieltjes integral over S of any continuous function $h(P)$ with respect to $\nu(e)$ or $\mu(e)$, and by means of the postulates (C), (A), (L), (M) of Daniell, generalize these integrals so that they apply to integrands $h(P)$ which are not necessarily continuous, but may in particular be merely bounded and measurable Borel.† We may then define the two functions of regular curves

$$\begin{aligned}\nu(w) &= \int_S q(P,w)\,d\nu(e_P),\\ \mu(w) &= \int_S q(P,w)\,d\mu(e_P),\end{aligned} \tag{3.1}$$

where $q(P,w)$ is the symmetric surface density at P of the region ω bounded by w.‡ The functions $\nu(w)$ and $\mu(w)$ are additive and bounded functions of regular curves on S.

---

* Bray and Evans, "A Class of Functions Harmonic within the Sphere," *American Journal of Mathematics*, Vol. 49 (1927), pp. 153-180. The definitions of lattice there given, on pp. 156 and 169 should be completed by making each lattice a partition of the previous one.

† P. J. Daniell, "A General Form of Integral," *Annals of Mathematics*, Vol. 19 (1918), pp. 279-294.

‡ This density may be defined as $\lim\limits_{\rho\to 0} (\sigma_\rho \cdot \omega')/\pi\rho^2$, where $\sigma_\rho \cdot \omega'$ denotes the area common to the projection of ω on the plane tangent to the surface at P and the circle σ_ρ in that plane of center P and radius ρ. Thus if P is an interior point of ω, $q(P,w)=1$, if an exterior point, $q(P,w)=0$, and if a point on the boundary w, $q(P,w)=\Psi/2\pi$ where Ψ is the angle from the forward to the backward tangent to w at P.

Moreover we may form Stieltjes integrals of the form

$$\int_S h(P)\,d\nu(w_P), \tag{3.2}$$

where $h(P)$ is continuous, and evaluate the integral in terms of a Riemann sum based on arbitrary modes of division of S into regular cells of diameter $< \delta$, where $\lim \delta = 0$, and the integral will be identical with $\int h(P)\,d\nu(e_P)$.* The integral (3.2) may also be extended to functions $h(P)$ bounded and measurable Borel, and in particular to $h(P) = q(P, w)$. Since then

$$\int_S q(P, w)\,d\nu(e_P) = \int_S q(P, w)\,d\nu(w_P)$$

with a similar identity for the function μ, we have

$$\begin{cases} \nu(w) = \displaystyle\int_S q(P, w)\,d\nu(w_P) \\ \mu(w) = \displaystyle\int_S q(P, w)\,d\mu(w_P). \end{cases} \tag{3.3}$$

Functions of curves which satisfy (3.3) are said to have *regular* discontinuities.† The functions defined by (3.1) are then not only bounded and additive, but have regular discontinuities.

In particular, therefore, we may rewrite the integrals (1) and (2) in the forms

$$v(M) = \int_S (1/MP)\,d\mu(w_P), \tag{3.4}$$

$$u(M) = \int_S \frac{\cos(MP, n_P)}{MP^2}\,d\nu(w_P), \tag{3.5}$$

where the $\mu(w)$ and $\nu(w)$ satisfy (3.3). We need also to consider the integral

(3.6)

$$v'(Q) = \int_S \frac{\cos(QP, n_Q)}{QP^2}\,d\mu(w_P)$$

$$= \int_S \frac{\cos(QP, n_Q)}{QP^2}\,d\mu(e_P)$$

where Q is on S.

---

* Bray and Evans, *loc. cit.*, p. 159.

† In analogy with regular discontinuities of functions of a single variable. See Bray and Evans, *loc. cit.*, pp. 157, 170. Stieltjes integrals may also be formed with respect to functions of curves whose discontinuities are not regular, and evaluated by means of an arbitrary mode of division. See R. N. Haskell, "A Note on Stieltjes Integrals," *Annals of Mathematics*, Vol. 29 (1928), pp. 543-548.

When M is not on S these integrals offer no difficulty. But (3.6), and (3.4) and (3.5), when $M = Q$ on S need special treatment. Consider the case of (3.5). In this case the integral is not defined as the limit of a sum, but is regarded as a generalized or improper integral.* That this integral exists for almost all Q on S and represents a summable function on S, and that the identity

$$(3.7)\qquad \int_\omega d\omega_Q \int_S \frac{\cos(QP, n_P)}{QP^2}\, dv(w_P) = \int_S dv(w_P) \int_\omega \frac{\cos(QP, n_P)}{QP^2}\, d\omega_Q$$

is valid, follow from the fact that the generalized integral

$$\int_\omega dt(w_P) \int_S \frac{|\cos(QP, n_P)|}{QP^2}\, d\omega_Q$$

is convergent, where $t(w)$ is a bounded additive function of positive type, with regular discontinuities, say the total variation function of $v(w)$.† Similarly,

$$(3.8)\qquad \int_\omega d\omega_Q \int_S \frac{\cos(QP, n_Q)}{QP^2}\, d\mu(w_P) = \int_S d\mu(w_P) \int_\omega \frac{\cos(QP, n_Q)}{QP^2}\, d\omega_Q$$

is valid. These results may be summarized in the following theorem.

THEOREM 1. *The integrals* (3.4), (3.5), *when M is a point Q on S, and the integral* (3.6) *converge for almost all Q, and the identities* (3.7), (3.8) *are valid. In these formulae $v(w)$ and $\mu(w)$ are given in terms of $v(e)$ and $\mu(e)$ by* (3.1), *and have regular discontinuities, but the analogous formulae may be written directly in terms of $v(e)$ and $\mu(e)$.*

4. *Boundary values of integrals.* With reference to the correspondence between points of S and points of S' with $|n(P)| < \tau$, as described in § 2, let w be a regular closed curve on S, and $w' = w'(\tau)$, $\omega' = \omega'(\tau)$ be the corresponding sets of points on S'. We indicate that S' approaches S from the interior or from the exterior, respectively, by the symbols $\lim\limits_{\tau=0+}$ and $\lim\limits_{\tau=0-}$ and consider the quantities $\lim\limits_{\tau=0+} U(\tau, w)$, $\lim\limits_{\tau=0-} U(\tau, w)$, $\lim\limits_{\tau=0+} V(\tau, w)$, $\lim\limits_{\tau=0-} V(\tau, w)$, where

$$(4.1)\qquad \begin{aligned} U(\tau, w) &= \int_{\omega'} u(M)\, d\omega_M, \\ V(\tau, w) &= \int_{\omega'} \frac{dv(M)}{dn_M}\, d\omega_M, \end{aligned}$$

* Daniell, *loc. cit.* Also Evans, "Fundamental Points of Potential Theory," *Rice Institute Pamphlet*, Vol. 7 (1920), pp. 252-329, in particular pp. 257-260.

† Evans, *loc. cit.*, p. 258. That the total variation function has regular discontinuities is proved as in Bray and Evans, *loc. cit.*, p. 170.

the integrations being extended over the surface S', and n_M being the normal to S' at M. Of course the quantity $V(\tau, w)$ depends merely on the curve w' and not on the particular surface on which it happens to lie, and represents merely the flux of $v(M)$ through that curve.

THEOREM 2. *As S' approaches S entirely from the inside or entirely from the outside of S the following equations hold:*

$$(4.2) \quad \lim_{\tau=0\pm} U(\tau, w) = \mp 2\pi v(w) + \int_S dv(e_P) \int_{\omega} \frac{\cos(QP, n_P)}{QP^2} d\omega_Q,$$

$$(4.3) \quad \lim_{\tau=0\pm} V(\tau, w) = \mp 2\pi \mu(w) + \int_S d\mu(e_P) \int_{\omega} \frac{\cos(QP, n_Q)}{QP^2} d\omega_Q.$$

Let $\omega'(\tau, \delta, P)$ be the portion of $\omega'(\tau)$ which is cut out by the normals to S in a small neighborhood $\omega(\delta, P)$ of P, where the projection of $\omega(\delta, P)$ on the tangent plane at P is a circle with center P and radius δ. Indicate the respective boundaries of these regions by $w'(\tau, \delta, P)$ and $w(\delta, P)$, and their complementary regions in $\omega'(\tau)$ and ω respectively by $\Omega'(\tau, \delta, P)$ and $\Omega(\delta, P)$.

With P on S and M on S', let ϕ be the angle (MP, n_P), and ϕ' the angle (MP, n_M), n_M being the normal to S' at M. If δ is sufficiently small, we shall have, for M in $\omega'(\tau, \delta, P)$,

$$\frac{|\cos\phi - \cos\phi'|}{r^2} \leq \frac{2}{r^2} \sin \frac{|\phi - \phi'|}{2} \leq \frac{\sphericalangle(n_M, n_P)}{r^2}$$
$$< Cr^{\alpha-2} < C\rho^{\alpha-2},$$
$$d\omega_M < 2\rho d\rho d\theta,$$

where ρ, the projection of $r = MP$, and θ are polar coördinates in the tangent plane at P. Hence

$$(4.4) \qquad \int_{\omega'(\tau,\delta,P)} \frac{|\cos\phi - \cos\phi'|}{r^2} d\omega_M < \frac{4\pi C\delta^\alpha}{\alpha}.$$

Hence, given ϵ, we may take $\delta(\epsilon)$, and then $\tau(\delta)$ so small that, by (4.4),

$$\left| \int_{\omega'(\tau,\delta,P)} \frac{\cos\phi}{r^2} d\omega_M \pm 2\pi q(P, w) \right| < \epsilon,$$

$q(P, w)$ being the density function, and the $+$ or $-$ sign being used according as S' is interior or exterior to S. In fact,

$$\int_{\omega'(\tau,\delta,P)} \frac{\cos\phi'}{r^2} d\omega_M$$

is the measure of the solid angle subtended by $\omega'(\tau, \delta, P)$ at P, prefixed by

the algebraic sign opposite to that of τ, and on account of the inequalities given for z in (2.5) and (2.9) differs from $2\pi q(P, w)$ by less than $4\pi C\rho$ as τ approaches zero.

Hence

$$(4.5) \qquad \left| \int_S d\nu(e_P) \int_{\omega'(\tau,\delta,P)} \frac{\cos\phi}{r^2}\, d\omega_M \pm 2\pi \int_S q(P, w)\, d\nu(e_P) \right| < \epsilon N(S),$$

where $N(S)$ is the total absolute mass on S.

On the other hand, given δ, the quantity $\int_{\Omega'(\tau,\delta,P)} \cos\phi/r^2 d\omega_M$ is continuous as τ approaches and becomes zero. Hence, given ϵ' and δ, we may take τ so small that

$$(4.6) \qquad \left| \int_S d\nu(e_P) \int_{\Omega'(\tau,\delta,P)} \frac{\cos\phi}{r^2}\, d\omega_M - \int_S d\nu(e_P) \int_{\Omega(\delta,P)} \frac{\cos\phi}{r^2}\, d\omega_Q \right| < \epsilon'.$$

But by (2.5) and (2.9) we have $|\cos(QP, n_P)| < K\rho$ uniformly in P for Q on S in the neighborhood of P. Consequently

$$\left| \int_{\omega(\delta,P)} \frac{\cos\phi}{r^2}\, d\omega_Q \right| \leq 4\pi K\delta,$$

and

$$(4.7) \qquad \left| \int_S d\nu(e_P) \int_{\Omega(\delta,P)} \frac{\cos\phi}{r^2}\, d\omega_Q - \int_S d\nu(e_P) \int_\omega \frac{\cos\phi}{r^2}\, d\omega_Q \right| < \epsilon'$$

if $\delta \leq \epsilon'/[4\pi K N(S)]$.

Hence, given $\bar{\epsilon}$, we can, by combining the inequalities (4.5), (4.6) and (4.7), choose δ small enough and then τ small enough so that

$$\left| \int_S d\nu(e_P) \int_{\omega'(\tau)} \frac{\cos\phi}{r^2}\, d\omega_M - \int_S d\nu(e_P) \int_\omega \frac{\cos\phi}{r^2}\, d\omega_Q \pm 2\pi\nu(w) \right| < \bar{\epsilon},$$

since $\nu(w) = \int_S q(P, w)\, d\nu(e_P)$. We have then finally

$$\lim_{\tau=0\pm} \int_S d\nu(e_P) \int_{\omega'(\tau)} \frac{\cos\phi}{r^2}\, d\omega_M = \mp 2\pi\nu(w) + \int_S d\nu(e_P) \int_\omega \frac{\cos\phi}{r^2}\, d\omega_Q,$$

which is the equation (4.2), which was to be proved.

Equation (4.3) is established in a similar way. We have

$$(4.8) \qquad V(\tau, w) = \int_{\omega'} d\omega_M \int_S \frac{\cos(MP, n_M)}{MP^2}\, d\mu(e_P)$$

$$= \int_S d\mu(e_P) \int_{\omega'} \frac{\cos\phi'}{r^2}\, d\omega_M$$

and if we write $\omega' = \omega'(\tau, \delta, P) + \Omega'(\tau, \delta, P)$, the proof proceeds as before.

Moreover, if n_Q denotes the normal to S at the point Q which, on S, corresponds to M on S' and we form $V_1(\tau, w)$:

$$V_1(\tau, w) = \int_{\omega'(\tau)} \frac{dv(M)}{dn_Q}\, d\omega_M \tag{4.9}$$

$$= \int_S d\mu(e_P) \int_{\omega'} \frac{\cos(MP, n_Q)}{MP^2}\, d\omega_M$$

and again write $\omega' = \omega'(\tau, \delta, P) + \Omega'(\tau, \delta, P)$, the quantity

$$\int_{\Omega'(\tau,\delta,P)} [\cos(MP, n_Q)/MP^2]\, d\omega_M$$

is continuous as τ approaches and becomes zero. But, denoting $\sphericalangle(MP, n_Q)$ by ϕ_1, we have for M in $\omega'(\tau, \delta, P)$

$$|\cos\phi_1 - \cos\phi'| \leq 2 \left| \sin\frac{\phi_1 - \phi'}{2} \right| \leq |\phi_1 - \phi'| \leq |\sphericalangle(n_Q, n_M)|$$

$$< C\overline{MQ}^a \leq C\overline{MP}^a$$

by (2.10), if δ has been chosen small enough. Hence we may substitute ϕ' for ϕ_1, as we substituted ϕ' for ϕ in the consideration of $U(\tau, w)$, and we have the following corollary.

COROLLARY. *The function $V_1(\tau, w)$ behaves in the same way as $V(\tau, w)$ when τ approaches zero,* i. e.,

$$\lim_{\tau=0\pm} V_1(\tau, w) = \lim_{\tau=0\pm} V(\tau, w).$$

5. *Generalized boundary value problems, and Stieltjes integral equations.* Let $F(w)$, $G(w)$ be arbitrary bounded additive functions of regular curves on S, with regular discontinuities. We speak of functions, given by potentials of a single layer (1), as of class (1), and of those given by potentials of a double layer (2), as of class (2).

THEOREM 3. *There is one and only one function $u(M)$ of class (2) for which*

$$\frac{1+\lambda}{2\lambda} U(0+, w) - \frac{1-\lambda}{2\lambda} U(0-, w) = F(w), \qquad (\lambda \neq 0), \tag{5.1}$$

and there is one and only one $v(M)$ of class (1) for which

$$\frac{1+\lambda}{2\lambda} V(0-, w) - \frac{1-\lambda}{2\lambda} V(0+, w) = G(w), \qquad (\lambda \neq 0) \tag{5.2}$$

unless λ *is a characteristic value of the kernel*

$$\frac{\lambda}{2\pi}\;\frac{\cos(QP, n_P)}{QP^2}\,.$$

THEOREM 4. *The value* $\lambda = +1$ *is not a characteristic value. Hence the solutions of the particular problems*

$$U(0+, w) = F(w), \qquad V(0-, w) = G(w)$$

are unique in the respective classes (2) *and* (1).

THEOREM 5. *The value* $\lambda = -1$ *is a characteristic value. For this value of* λ, *condition* (5.1) *can be satisfied if and only if*

$$\int_S \phi_2(P)\,dF(e_P) = 0,$$

where $\phi_2(P)$ *is a solution of the homogeneous equation with kernel*

$$\frac{\lambda}{2\pi}\;\frac{\cos(QP, n_Q)}{QP^2}\,, \qquad (\lambda = -1),$$

and is determined except for an arbitrary multiplicative constant.

The mass function $\nu(w)$ *of* (2) *is then determined except for an arbitrary additive term of the form* $C\omega$, *which does not change the value of* $u(M)$, M *exterior to* S.

Condition (5.2) *can be satisfied for* $\lambda = -1$, *if and only if*

$$G(S) = 0.$$

The mass function $\mu(w)$ *of* (1) *is then determined except for an arbitrary additive term of the form* $C\int_\omega \phi_2(P)\,d\omega_P$; *the corresponding* $v(M)$ *contains an arbitrary additive constant if* M *is interior to* S.

The value $\lambda = -1$ in (5.1) corresponds to the exterior Dirichlet problem, in (5.2) to the interior Neumann problem.

On account of (4.2), (4.3), the conditions (5.1) and (5.2) may be rewritten as Stieltjes integral equations:

$$\nu(w) = -\frac{\lambda}{2\pi}F(w) + \frac{\lambda}{2\pi}\int_S d\nu(e_P)\int_\omega \frac{\cos(QP, n_Q)}{QP^2}\,d\omega_Q, \tag{5.3}$$

$$\mu(w) = \frac{\lambda}{2\pi}\,G(w) - \frac{\lambda}{2\pi}\int_S d\mu(e_P)\int_\omega \frac{\cos(QP, n_Q)}{QP^2}\,d\omega_Q. \tag{5.4}$$

If $v(M)$ is a function of class (1) which satisfies (5.2), the function $\mu(w)$, bounded and additive, with regular discontinuities, must be a solution of (5.4). Conversely, if $\mu(w)$ is a solution of (5.4), bounded and additive, and having regular discontinuities, the function $v(M)$ given by (1) will satisfy (5.2). Similarly, for the functions $u(M)$ of class (2).

Each of the equations (5.3), (5.4) has a unique solution which is a bounded additive function of curves, with regular discontinuities, unless λ is a characteristic value of the kernel

$$\frac{\lambda}{2\pi}\,\frac{\cos(QP, n_P)}{QP^2}.$$

Equations (5.3), (5.4) are solved by means of Fredholm equations, by means of the substitution

$$\nu(w)+\frac{\lambda}{2\pi}F(w)=R_1(w), \tag{5.5}$$

$$\mu(w)-\frac{\lambda}{2\pi}G(w)=R_2(w). \tag{5.6}$$

The possibility of reducing (5.3), (5.4) to Fredholm equations depends on the fact that $R_1(w)$ and $R_2(w)$ are absolutely continuous, on account of (3.7). Consider (5.3). Substituting the value of $\nu(w)$ from (5.5) into (5.3), we have

$$R_1(w)=-\frac{\lambda^2}{4\pi^2}\int_S dF(w_P)\int_\omega \frac{\cos(QP, n_P)}{QP^2}\,d\omega_Q + \frac{\lambda}{2\pi}\int_S dR_1(w_P)\int_\omega \frac{\cos(QP, n_P)}{QP^2}\,d\omega_Q.$$

The derivative nearly everywhere of $R_1(w)$ is

$$r_1(Q)=-\frac{\lambda^2}{4\pi^2}\int_S \frac{\cos(QP, n_P)}{QP^2}\,dF(w_P) + \frac{\lambda}{2\pi}\int_S \frac{\cos(QP, n_P)}{QP^2}\,dR_1(w_P),$$

or

$$r_1(Q)=f(Q)+\frac{\lambda}{2\pi}\int_S \frac{\cos(QP, n_P)}{QP^2}\,r_1(P)\,d\omega_P, \tag{5.7}$$

since $R_1(w)$ is absolutely continuous, where we have placed

$$f(Q)=-\frac{\lambda^2}{4\pi^2}\int_S \frac{\cos(QP, n_P)}{QP^2}\,dF(w_P).$$

In the same way, we have

$$(5.8) \qquad r_2(Q) = g(Q) - \frac{\lambda}{2\pi}\int_S \frac{\cos(QP, n_Q)}{QP^2} r_2(P)\,d\omega_P,$$

where

$$g(Q) = -\frac{\lambda^2}{4\pi^2}\int_S \frac{\cos(QP, n_Q)}{QP^2}\,dG(w_P).$$

Hence if $\nu(w)$ is a solution of (5.3), bounded and additive, with regular discontinuities, $r_1(Q)$ will be a solution of (5.7). That is, (5.7) is necessary for (5.3); and (5.4) implies (5.8). Conversely, if $r_1(Q)$ is a summable solution of (5.7), then the $\nu(w)$ given by (5.5) will be bounded and additive, and will satisfy (5.3). Similar statements may be made with regard to (5.8), (5.6) and (5.4).

LEMMA. *Unless λ is a characteristic value,* i.e., *unless for this value of λ there is a continuous solution of the homogeneous equation with kernel*

$$\frac{\lambda}{2\pi}\;\frac{\cos(QP, n_P)}{QP^2},$$

each of the equations (5.7), (5.8) *has a unique summable solution.*

Suppose λ not to be a characteristic value. Then (5.7) has a summable solution. In fact, this equation is equivalent to that which is obtained by twice repeated substitution of (5.7) into itself, the resulting equation having a bounded continuous kernel.* Since the known function in this equation is summable, the solution, given in terms of the resolvent kernel is summable. The solution moreover is unique. For, suppose there were two different summable solutions of (5.7). Then their difference, say $\phi_1(Q)$, would be a summable solution of the homogeneous equation

$$(5.9) \qquad \phi_1(Q) = \frac{\lambda}{2\pi}\int_S \frac{\cos(QP, n_P)}{QP^2}\,\phi_1(P)\,d\omega_P.$$

But every solution of this equation is a solution of the homogeneous equation with twice iterated (and continuous) kernel, namely, the equation

$$(5.10) \qquad \phi_1(Q) = \frac{\lambda^3}{8\pi^3}\int_S K_3(Q,P)\,\phi_1(P)\,d\omega_P.$$

If $\phi_1(Q)$ is summable, it is continuous in virtue of (5.10). But this means

* Goursat, *Cours d'Analyse*, Vol. 3, Paris (1915), pp. 355, 364; Kellogg, *Foundations of Potential Theory*, Berlin (1929), p. 301. Goursat shows merely that the twice iterated kernel is bounded, and Kellogg, that it is continuous on the basis of proper assumptions on S. The assumptions on S in the proof cited are slightly more restrictive than (2.9), but the proof remains valid merely under (2.9).

that (5. 9) has a continuous non-zero solution. Consequently λ would be a characteristic value, contrary to the assumption. The lemma is proved.

Since the kernels

$$\frac{\lambda}{2\pi}\frac{\cos(QP, n_P)}{QP^2}, \qquad -\frac{\lambda}{2\pi}\frac{\cos(QP, n_Q)}{QP^2}$$

of (5. 7) and (5. 8) respectively are associated, and have the same characteristic values, we may consider both equations at the same time. It is known that $\lambda = -1$ is a characteristic value for these kernels, while $\lambda = 1$ is not.*

* These statements are usually proved on the basis of slightly more restrictive assumptions on S, as in Kellogg, *loc. cit.*, p. 311. Accordingly we indicate a brief proof on the basis of our hypotheses with regard to S, of which the characteristic assumption is (2. 9).

Suppose $\lambda = 1$ were a characteristic value. There would then be a continuous, not identically vanishing solution $\phi(Q)$, of the homogeneous integral equation corresponding to (5. 8), namely

$$(\alpha) \qquad 0 = \phi(Q) + \frac{\lambda}{2\pi}\int_S \frac{\cos(QP, n_Q)}{QP^2}\,\phi(P)\,d\omega_P, \qquad (\lambda = 1).$$

The function $v_0(M)$, given in the infinite region exterior to S as a potential of a single layer (1) on S,

$$v_0(M) = \int_S \frac{1}{MP}\,\phi(P)\,d\omega_P,$$

would then satisfy the boundary condition (4. 3) with

$$\mu(w) = \int_\omega \phi(P)\,d\omega_P,$$

that is,

$$\lim_{\tau=0-} V_0(\tau, w) = \int_\omega \{2\pi\phi(Q) + \int_S \frac{\cos(QP, n_Q)}{QP^2}\,\phi(P)\,d\omega_P\}\,d\omega_Q = 0.$$

If now we apply Green's theorem to the infinite region T', exterior to S', we have

$$\int_{T'}\{(\partial v_0/\partial x)^2 + (\partial v_0/\partial y)^2 + (\partial v_0/\partial z)^2\}\,dT = \int_{S'} v_0(\partial v_0/\partial n)\,d\omega'$$
$$= \int_{S'} v_0(M)\,dV_0(\tau, \omega_M),$$

since $v_0(M)$ vanishes canonically at ∞. But since $v_0(M)$ is a continuous function of Q and τ as M approaches Q along n_Q, and the total variation of $V_0(\tau, \omega)$ is bounded, we have by a well known theorem

$$(\beta) \qquad \lim_{\tau=0-}\int_{S'} v_0(M)\,dV_0(\tau, \omega_M) = \int_{S'} v_0(M)\,d[\lim_{\tau=0-} V_0(\tau, \omega)] = 0.$$

In fact, the total variation of $V_0(\tau, \omega)$ does not exceed the quantity

$$\int_S d\omega_P\,|\,\phi(P)\,|\int_{S'}\left|\frac{\cos(MP, n_M)}{MP^2}\right|\,d\omega_M,$$

As to $\lambda = -1$, (5.9) has obviously the solution $r_1(Q) = 1$. Hence the homogeneous equation corresponding to (5.8) has a not identically vanishing solution $\phi_2(Q)$. Both of these homogeneous equations have the same number of linearly independent solutions. It is known that for $\lambda = -1$ there cannot be two linearly independent solutions of (5.9) or of the homogeneous equation corresponding to (5.8).*

For the value $\lambda = -1$, a necessary and sufficient condition that (5.7) have a solution is that

$$\int_S \phi_2(P) f(P)\, dP = 0$$

which, on replacing $f(P)$ by its equivalent in terms of $F(w)$, and reversing the order of integration, takes the form

$$(5.11) \qquad \int_S \phi_2(P)\, dF(w_P) = 0,$$

$\phi_2(P)$ denoting a solution of the homogeneous equation with kernel

---

and if we write

$$\int_{S'} = \int_{\omega'(\tau,\delta,P)} + \int_{\Omega'(\tau,\delta,P)},$$

the first of these integrals is $< 4\pi$ and the second is $<$(const.) (meas. S') $<$ (const.) $\times$ (2 meas. S).

Hence if T is the region exterior to S, we have, from (β),

$$\int_T \{(\partial v_0/\partial x)^2 + (\partial v_0/\partial y)^2 + (\partial v_0/\partial z)^2\} dT = 0.$$

But this means that v_0 is constant outside S, and since it vanishes at ∞ and is continuous across S, it vanishes outside and on S. Hence it vanishes identically also inside S. The condition (4.3) yields now the fact that

$$\int_\omega \{-2\pi\phi(Q) + \int_S \frac{\cos(QP, n_Q)}{QP^2}\, \phi(P)\, d\omega_P\} d\omega_Q = 0,$$

which with (α) and the continuity of $\phi(Q)$ implies that $\phi(Q) \equiv 0$. This is contrary to the hypothesis. Hence $\lambda = 1$ is not a characteristic value.

* Suppose in fact that we apply Green's theorem to the region interior to S, in the same way as we applied it in the previous footnote to the region exterior to S, taking as $v(M)$ the potential of a single layer associated with a continuous solution $\phi_2(Q)$. This potential $v(M)$ will be constant within and on S and will vanish canonically at ∞. The potentials defined by two such solutions $\phi_2(Q)$ will therefore be linearly dependent in all space. The functions $\phi_2(Q)$ will also therefore be linearly dependent, by means of the relation

$$\lim_{\tau=0+} V(\tau, w) - \lim_{\tau=0-} V(\tau, w) = -4\pi \int_\omega \phi_2(P)\, d\omega_P,$$

which is a consequence of (4.3).

$$-\frac{\lambda}{2\pi}\,\frac{\cos(QP, n_Q)}{QP^2}, \qquad (\lambda = -1).$$

Suppose the condition (5. 11) is satisfied. Then the general solution of (5. 7) is

$$r_1(Q) = \bar{r}_1(Q) + \bar{C},$$

since $\phi_1(Q) = \bar{C}$ is a solution of the homogeneous equation

$$\phi_1(Q) = -\frac{1}{2\pi}\int_S \frac{\cos(QP, n_P)}{QP^2}\,\phi_1(P)\,d\omega_P.$$

Hence by (5. 5), we have

$$\begin{aligned} \nu(w) &= \frac{1}{2\pi}F(w) + R_1(w) \\ &= \frac{1}{2\pi}F(w) + \int_\omega \bar{r}_1(Q)\,d\omega_Q + \int_\omega \bar{C}\,d\omega_Q \\ &= \nu_1(w) + \bar{C}\omega. \end{aligned}$$

If we substitute this value of $\nu(w)$ in (2), dealing as we are here with the exterior Dirichlet problem, the contribution due to the term $\bar{C}\omega$ is zero, and $u(M)$ is given simply by the formula (2) with $\nu(e) = \nu_1(e)$.

For the value $\lambda = -1$, and the interior Neumann problem, a necessary and sufficient condition that (5. 8) have a solution is that

$$\int_S g(Q)\phi_1(Q)\,d\omega_Q = \bar{C}\int_S g(Q)\,d\omega_Q = 0.$$

But this is equivalent to the condition

$$0 = -\frac{1}{4\pi^2}\int_S dG(w_P)\int_S \frac{\cos(QP, n_Q)}{QP^2}\,d\omega_Q = \frac{1}{2\pi}\int_S dG(w) = G(S).$$

Hence we have the condition

(5. 12) $$G(S) = 0.$$

But, since the steps may be retraced, this condition also is both necessary and sufficient in order that the interior Neumann problem be solved by means of a potential of a single layer.

If (5. 12) is satisfied, the general solution of (5. 8) is in the form

$$r_2(Q) = \bar{r}_2(Q) + C\phi_2(Q)$$

where C is an arbitrary constant, and $\phi_2(Q)$ is a solution of the homogeneous equation

$$\phi_2(Q) = \frac{1}{2\pi}\int_S \frac{\cos(QP, n_Q)}{QP^2}\,\phi_2(P)\,d\omega_P,$$

and is a continuous function. But as we have seen [in the preceding footnote] the potential (1) where $\nu(w) = \int_\omega \phi_2(P)\,d\omega_P$ reduces inside S to a constant. Hence the effect of the term $C\phi_2(Q)$ is merely to change $v(M)$ inside S by an arbitrary constant.

6. *Boundary values of $u(M)$ and $v(M)$.* For simplicity, we consider merely approach in the narrow sense, letting M approach Q on S along the normal n_Q ($\lim M = Q\pm$). We shall show that $u(M)$ and $dv(M)/dn_Q$ approach definite boundary values wherever $\nu(w)$ and $\mu(w)$ have derivatives, that is, nearly everywhere on S; and for S we may take a surface which satisfies merely the conditions of the Lemma of § 2, with no assumptions about the curvature. We say that $\nu(w)$ [or $\nu(e)$] has the derivative A at Q if $\lim\limits_{\omega=0} \nu(w)/\omega = A$, where the regions ω form a regular family about Q.

THEOREM 6. *Let S satisfy the conditions of the Lemma of § 2, and let Q be a point on S where $\nu(w)$ has a derivative $\nu'(Q) = A$; then*

$$\lim_{M=Q\pm} u(M) = \mp 2\pi A + \int_S \frac{\cos(QP, n_P)}{QP^2}\, d\nu(e_P). \tag{6.1}$$

If $\mu'(Q)$ exists, $= A$, then

$$\lim_{M=Q\pm} \frac{dv(M)}{dn_Q} = \mp 2\pi A + \int_S \frac{\cos(QP, n_Q)}{QP^2}\, d\mu(e_P). \tag{6.2}$$

So far, in dealing with integrals over the whole of S it has made no difference whether we used the differential $d\nu(e)$ or the differential $d\nu(w)$. For integrals over a part of S, however, the symbols may involve some ambiguity. Hence for the integral $\int_w f(P)\,d\nu(w_P)$, extended over a region ω bounded by w, where $f(P)$ is integrable in the Borel sense, we introduce the definition

$$\int_w f(P)\,d\nu(w_P) = \int_S q(P, w) f(P)\,d\nu(w_P) = \int_S q(P, w) f(P)\,d\nu(e_P),$$

in which $q(P, w)$ is the density function for ω bounded by w. The integral is thus an additive function of curves w.

We indicate integration over the region Ω complementary to ω by the symbol $\int_W f(P)\,d\nu(w_P)$, W being the same curve as w, regarded, if we like, as having opposite sense. We have

$$\int_w f(P)\,d\nu(w_P) + \int_W f(P)\,d\nu(w_P) = \int_S f(P)\,d\nu(w_P).$$

It may be noted that given $\nu(w)$ there is one and only one $\nu(e)$ such that $\nu(w)=\nu(\omega)$ for curves w where $\nu(w)$ is continuous in the sense of Volterra; conversely, given $\nu(e)$ there is one and only one $\nu(w)$ corresponding to it in this sense, provided $\nu(w)$ has regular discontinuities.*

Since $\nu(w)$ and $\mu(w)$ are independent, we may as well assume that the hypotheses for each of them are satisfied at some point Q, and prove both identities at once. We must first examine the integral

$$I=\int_S \frac{\cos(QP, n_P)}{QP^2}\, d\nu(w_P)=\int_S \frac{\cos(QP, n_P)}{QP^2}\cdot d\nu(e_P)$$

and the similar integral J, with n_Q instead of n_P. We may write

$$\nu(w)=A\omega+\omega\eta(w)$$

where $\omega\eta(w)$ is a bounded additive function of regular curves on S, with regular discontinuities, such that $\lim\limits_{\omega\to 0}\eta(w)=0$ when the regions ω form a regular family about Q. Without loss of generality we may assume that $\eta(w)$ is of positive type.†

Let I_1 be that part of I referring to $A\omega$, and I_2 to $\omega\eta(w)$; the convergence of I_1 is a consequence of the fact that S is of class Γ. The convergence of I_2 will be established as a generalized integral, if it can be done when $\cos(QP, n_P)$ is replaced by $|\cos(QP, n_P)|$, and the convergence of the latter integral will be established if

$$\lim_{m\to\infty}\int_S h(m, P)\, d[\omega\eta(w)]$$

exists, where

$$\begin{aligned} h(m, P) &= \frac{|\cos(QP, n_P)|}{QP^2}, && (QP > 1/m),\\ &= m^2\,|\cos(QP, n_P)|, && \text{otherwise.}\end{aligned}$$

In fact the $h(m,P)$ form an increasing sequence of functions of P, with $\lim\limits_{m\to\infty} h(m, P)=|\cos(QP, n_P)|/QP^2$. Finally we need only consider the neighborhood $\omega(\delta, Q)$ of Q, so that $|\cos(QP, n_P)|<3g(\rho)$, ρ being the projection of QP on the tangent plane at Q, and $g(\rho)$ being the continuous function of § 2 such that $m(\delta)=\int_0^\delta [g(\rho)/\rho]\, d\rho$ converges. In fact,

* Bray and Evans, *loc. cit.*, p. 159.

† Bray and Evans, *loc. cit.*, p. 173. In the analysis there given, the case where one or the other of the two families p_i', p_j'' may contain merely a finite number of elements should be considered; but in this special case the proof is immediate.

$$\begin{aligned}&|\cos(QP, n_P)-\cos(QP, n_Q)|\\ &\qquad \leq 2\,|\sin\{(QP, n_P)+(QP, n_Q)\}/2|\;|\sin(n_P, n_Q)/2| \leq g(\rho).\end{aligned}$$

We may therefore substitute for the $h(m, P)$ the dominating increasing sequence

$$\begin{aligned}h_1(m, P)&=3g(\rho)/\rho^2 &&\text{if } \rho>1/m,\\ &=3m^2g(1/m) &&\text{otherwise.}\end{aligned}$$

The function $\omega\eta(w)$ being of positive type is a non-decreasing function of ρ if $w=w(\rho, Q)$. If we represent this function by $\beta(\rho)$ and remember that $h_1(m, P)$ is a function of ρ alone, we have

$$\int_{w(\delta,Q)} h_1(m, P)\,d[\omega\eta(w)]=\int_0^\delta h_1(m, P)\,d\beta(\rho)$$

and this by an integration by parts is equal in absolute value to

$$\begin{aligned}&\left|\,[\beta(\rho)h_1(m, P)]_0^\delta-\int_0^\delta \beta(\rho)\,dh_1(m, P)\,\right|\\ &\leq 3\left\{\frac{\beta(\delta)g(\delta)}{\delta^2}+\left|\int_{1/m}^\delta \beta(\rho)\,d\,\frac{g(\rho)}{\rho^2}\right|\right\}\\ &\leq 3g(\delta)\eta[w(\delta, Q)]+3\pi\eta[w(\delta, Q)]\left|\int_{1/m}^\delta \rho^2 d\,\frac{g(\rho)}{\rho^2}\right|\end{aligned}$$

The Stieltjes integral in the last expression is however, by a second integration by parts, the same as

$$\left|\,g(\delta)-g(1/m)-2\int_{1/m}^\delta \frac{g(\rho)}{\rho}\,d\rho\,\right|<g(\delta)+2m(\delta).$$

Hence

$$\int_{w(\delta,Q)} h_1(m, P)\,d[\omega\eta(w)]<6\pi\eta(w(\delta, Q))[g(\delta)+m(\delta)].$$

But this bound is independent of m, so that the point is proved. Similar considerations apply to the integral J. Incidentally, this bound may be made as small as we please by taking δ small enough.

It is well to emphasize the following results, implied by the analysis which has just been completed.

LEMMA. *The integrals*

$$\int_{\omega(\delta,Q)} [g(\rho)/\rho^2]\,d\omega,\quad \int_{w(\delta,Q)} [g(\rho)/\rho^2]\,d\,[\omega\eta(w)]$$

are convergent, and may be made arbitrarily small with δ, *if* $g(\rho)$ *is a positive continuous increasing function of* ρ *such that* $\int_0^\delta [g(\rho)/\rho]\,d\rho$ *is convergent.*

18

COROLLARY. *We have, under the hypotheses of Theorem 6,*

$$(6.3)\qquad \lim_{\delta=0}\int_{\Omega(\delta,Q)}\frac{\cos(QP,n_P)}{QP^2}\,d\omega=\int_S\frac{\cos(QP,n_P)}{QP^2}\,d\omega,$$

$$(6.4)\qquad \lim_{\delta=0}\int_{W(\delta,Q)}\frac{\cos(QP,n_P)}{QP^2}\,d\nu(w_P)=\int_S\frac{\cos(QP,n_P)}{QP^2}\,d\nu(w_P),$$

with similar relations if n_P is replaced by n_Q and $\nu(w)$ by $\mu(w)$.

If we let ϕ be the angle (MP,n_P), ϕ' the angle (MP,n_Q), $\psi=\phi-\phi'$, $r=MP$, we have

$$(6.5)\qquad u(M)=\int_S\frac{\cos\phi}{r^2}\,d\nu(w_P),\qquad \frac{d\nu(M)}{dn_Q}=\int_S\frac{\cos\phi'}{r^2}\,d\mu(w_P).$$

We evaluate the limits of these integrals first by comparing them for the portion $\omega(\delta,Q)$ of S, then with corresponding integrals over the projection of $\omega(\delta,Q)$ on the tangent plane at Q, and finally by noticing that the integrals over the portion $\Omega(\delta,Q)$ are continuous as M approaches Q.

Accordingly, we show first that given ϵ we can choose δ small enough so that for all smaller values of δ we have

$$(6.6)\qquad \int_{\omega(\delta,Q)}\left|\frac{\cos\phi-\cos\phi'}{r^2}\right|d\omega\leq\epsilon$$

$$(6.7)\qquad \int_{w(\delta,Q)}\left|\frac{\cos\phi-\cos\phi'}{r^2}\right|d[\omega\eta(w)]\leq\epsilon,$$

independently of the position of M. We have in fact, if P is in $\omega(\delta,Q)$,

$$\left|\frac{\cos\phi-\cos\phi'}{r^2}\right|=\frac{2}{r^2}\left|\sin\frac{\phi+\phi'}{2}\sin\psi/2\right|\leq\psi/r^2\leq g(\rho)/\rho^2,$$

applying (2.4) and making use of the fact that $|\psi|\leq|n_P,n_Q|$. The desired inequalities follow then immediately from the lemma.

Let r_1 be MP_1 and ϕ_1 be the angle MP_1,n_Q where P_1 is the point, on the tangent plane at Q, at the foot of the perpendicular from P. We show now that, given ϵ, we can choose δ small enough so that for all smaller δ we have

$$(6.8)\qquad \int_{w(\delta,Q)}\left|\frac{\cos\phi'}{r^2}-\frac{\cos\phi_1}{r_1^2}\right|d([\omega\eta(\omega)]\leq\epsilon/2,$$

independently of the position of M. In fact, using the symbol z of § 2,

$$\left|\frac{\cos\phi'}{r^2}-\frac{\cos\phi_1}{r_1^2}\right|=\left|\frac{\cos\phi'-\cos\phi_1}{r^2}+\frac{\cos\phi_1(r_1+r)(r_1-r)}{r^2r_1^2}\right|$$

$$\leq 2g(\rho)/\rho^2+2z/\rho^3\leq 2g(\rho)/\rho^2+4g(\rho)/\rho^2=6g(\rho)/\rho^2,$$

the last inequality being a consequence of (2.5). Consequently, here again, the desired inequality follows immediately from the lemma.

Moreover if δ is small enough we can show that

$$\left| \int_{w(\delta,Q)} (\cos \phi_1/r_1^2) d[\omega\eta(w)] \right| < \epsilon/2,$$

whence, from (6.8),

(6.9) $$\left| \int_{w(\delta,Q)} (\cos \phi'/r^2) d[\omega\eta(w)] \right| \leq \epsilon,$$

independently of M.

For this purpose we introduce again the notation $\beta(\rho)$, and write, integrating by parts,

$$\left| \int \frac{\cos \phi_1}{r^2} d[\omega\eta(w)] \right| = \left| \int \frac{\cos \phi_1}{r_1^2} d\beta(\rho) \right|$$
$$\leq \beta(\delta)/\delta^2 + \left| \int \beta(\rho) d(\cos \phi_1/r_1^2) \right|,$$

of which the first term of the last member may be made as small as we please, with δ. The function $\cos \phi_1/r_1^2$ is a monotonic function of ρ, since for a given M the numerator decreases in numerical value, without changing sign, and the denominator increases, as ρ increases. In the second term, then, we introduce η_δ as the upper bound of $\eta(w)$ when w is the circle of radius ρ, $0 < \rho \leq \delta$, and obtain

$$\left| \int \beta(\rho) d(\cos \phi_1/r_1^2) \right| < 2\pi\eta_\delta \left| \int \rho^2 d(\cos \phi_1/r_1^2) \right|,$$

for, since $|(n_Q, n_P)| < 1$, $\omega(\rho) < 2\pi\rho^2$. But by performing again the integration by parts, the right hand member of this inequality is

$$\leq 2\pi\eta_\delta + 2\eta_\delta \left| \int (\cos \phi_1/r_1^2) d(\pi\rho^2) \right|$$
$$\leq 2\pi\eta_\delta + 4\pi\eta_\delta = 6\pi\eta_\delta,$$

and this also may be made as small as we please with δ, independently of M, since η_δ approaches zero with δ. Thus the desired inequality is established.

The rest of the proof is not difficult. In accordance with (6.4) we can take δ small enough so that, given ϵ,

(6.10) $$\left| \int_S \frac{\cos(QP, n_P)}{QP^2} d\nu(w_P) - \int_{W(\delta,Q)} \frac{\cos(QP, n_P)}{QP^2} d\nu(w_P) \right| \leq \epsilon.$$

Also, if we denote by γ the limit of the absolute value of the solid angle subtended at M by $\omega(\delta, Q)$ as M approaches Q, we can take δ small enough so that, given ϵ,

(6.11) $$|A\gamma - 2\pi A| \leq \epsilon.$$

So far we have dealt merely with inequalities depending on δ. When δ

is fixed, however, we may choose M close enough to Q on n_Q so that the following inequalities hold, when ϵ_1 is given:

$$(6.12)\qquad \left| \int_{W(\delta,Q)} \frac{\cos\phi}{r^2}\, d\nu(w_P) - \int_{W(\delta,Q)} \frac{\cos(QP, n_P)}{QP^2}\, d\nu(w_P) \right| \leq \epsilon_1,$$

$$(6.13)\qquad \left| \int_{\omega(\delta,P)} \frac{\cos\phi}{r^2}\, A\, d\omega \pm A\gamma \right| \leq \epsilon_1, \qquad M \text{ on } \pm n_Q.$$

Consider now the inequalities:

$$\begin{aligned}
R = &\left| \int_S \frac{\cos\phi}{r^2}\, d\nu(w_P) - \int_S \frac{\cos(QP, n_P)}{QP^2}\, d\nu(w_P) \pm 2\pi A \right| \\
\leq &\left| \int_S \frac{\cos(QP, n_P)}{QP^2}\, d\nu(w_P) - \int_{W(\delta,Q)} \frac{\cos(QP, n_P)}{QP^2}\, d\nu(w_P) \right| \\
+ &\left| \int_{w(\delta,Q)} \frac{\cos\phi}{r^2}\, d[\omega\eta(w)] - \int_{w(\delta,Q)} \frac{\cos\phi'}{r^2}\, d[\omega\eta(w)] \right| \\
+ &\left| \int_{w(\delta,Q)} \frac{\cos\phi'}{r^2}\, d[\omega\eta(w)] \right| \\
+ &\left| A\gamma - 2\pi A \right| \\
+ &\left| \int_{W(\delta,Q)} \frac{\cos\phi}{r^2}\, d\nu(w_P) - \int_{W(\delta,Q)} \frac{\cos(QP, n_P)}{QP^2}\, d\nu(w_P) \right| \\
+ &\left| \int_{\omega(\delta,Q)} \frac{\cos\phi}{r^2}\, A\, d\omega \pm A\gamma \right|
\end{aligned}$$

where with the notation $\pm$ the $+$ sign is taken if M is on $+n_Q$, the $-$ sign if M is on $-n_Q$. By taking δ small enough each of the first four expressions in absolute value signs, of the right hand member, may be made $\leq \epsilon$, by means of (6.10), (6.7), (6.9), (6.11), independently of the position of M; with δ so fixed, M may be taken on $\pm n_Q$ near enough to Q so that each of the last two expressions may be made $\leq \epsilon_1$, by means of (6.12), (6.13). Hence, given ϵ and ϵ_1 arbitrarily small, we can choose M near enough to Q so that

$$R \leq 4\epsilon + 2\epsilon_1.$$

But this establishes the equation (6.1) of the theorem.

By means of similar inequalities, taking account also of (6.6), we establish (6.2). This completes the demonstration of the theorem.

THE RICE INSTITUTE,
HOUSTON, TEXAS.

d-

?

n
as
y
;
of
).
Q

o-